NAVAL POSTGRADUATE SCHOOL

THESIS

A SURVEY AND SECURITY STRENGTH CLASSIFICATION OF PKI CERTIFICATE REVOCATION MANAGEMENT IMPLEMENTATIONS

by

John L. MacMichael, Jr.

December 2003

Thesis Advisor:	J. D. Fulp
Second Reader:	D. F. Warren

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 2003	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: A Survey and Security Strength Classification of PKI Certificate Revocation Management Implementations			**5. FUNDING NUMBERS**
6. AUTHOR(S) John L. MacMichael, Jr.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT ***(maximum 200 words)***

In this thesis, I define all currently operational, proposed, and theoretically possible methods of certificate revocation. The role of certificate revocation within the larger scheme of PKI is examined and the mandates upon Department of Defense from the Certification Practices Statement (CPS) and Certificate Policy (CP) are examined. A "best case" model for revocation is suggested. The security attributes affecting certificate revocation are examined; from these attributes a set of metrics are defined for the purpose of measuring the security-relevant strengths and weaknesses of all plausible methods of certificate revocation. Each method is examined and ranked according to security strength. Conclusions regarding certificate revocation use within Department of Defense are made and further study within the field is suggested.

14. SUBJECT TERMS PKI, X.509, OCSP, NOVOMODO, SCVP, CRL, Certificate Revocation, Security, DoD, Certificate Policy, Certification Practices Statement			**15. NUMBER OF PAGES** 101
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

A SURVEY AND SECURITY STRENGTH CLASSIFICATION OF PKI CERTIFICATE REVOCATION MANAGEMENT IMPLEMENTATIONS

John L. MacMichael, Jr.
Lieutenant Commander, United States Navy
B.S., Virginia Military Institute, 1988

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 2003

Author: John L. MacMichael, Jr.

Approved by: J. D. Fulp
Thesis Advisor

D. F. Warren
Second Reader

Dan C. Boger
Chairman, Department of Information Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

In this thesis, I define all currently operational, proposed, and theoretically possible methods of certificate revocation. The role of certificate revocation within the larger scheme of PKI is examined and the mandates upon Department of Defense from the Certification Practices Statement (CPS) and Certificate Policy (CP) are examined. A "best case" model for revocation is suggested. The security attributes affecting certificate revocation are examined; from these attributes a set of metrics are defined for the purpose of measuring the security-relevant strengths and weaknesses of all plausible methods of certificate revocation. Each method is examined and ranked according to security strength. Conclusions regarding certificate revocation use within Department of Defense are made and further study within the field is suggested.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

Many people deserve recognition for the guidance and tutelage they provided while I worked on this thesis. Foremost among them is J. D. Fulp; many of the original ideas in the document were the result sweating the details out with J.D. J.D. is a friend and mentor. I also owe a great debt to several persons in the industry who freely gave of their time: I was always amazed that these titans in their fields took time to respond to a graduate student's email. Each of these fine men took time to answer emails, phone calls, or take meetings: Dr. Carlisle Adams, Dr. Bob Van Spyk, Dr. William Flanigan, Dr John Hines; Dr. Andrew Nash, and Dr. Silvio Micali. Finally, I would like to thank my wife who regularly asked me, "When is that thing going to be done?" She remains the center of my universe and the impetus of my efforts.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

One of the impediments to broader use of Public Key Infrastructure (PKI) within the DoD is the non-standardized implementation of certificate revocation management within the growing field of PKI implementations. Certificate revocation plays a critical security role in any PKI as it provides a means of invalidating a certificate prior to that certificate's pre-determined expiration date. A PKI implementation that fails to verify the current validity of certificates it processes is susceptible to security policy violations. These violations could occur in many forms, but the simplest example is that of an employee who is fired prior to his certificate's normal expiration, and then subsequently signs or decrypts company documents using his company issued certificate. Due to the distributed nature of electronic data interchange and the latency inherent in updating certificate information in the issuing Certificate Authority's (CA) storage directory; management and use of the certificate revocation has important security policy ramifications. This thesis attempts to survey all methods of certificate revocation management so that DoD policy-makers can make informed decisions based upon the relative security strengths and interoperability considerations of each method. This paper spans both DoD and non-DoD implementations. It also spans all known current, proposed, and theoretically plausible certificate revocation implementations. The intention of this thesis is that the findings are inclusive of all implementations that may prove valid and prudent for DoD use or consideration.

B. STATEMENT OF PROBLEM

PKI has been on the verge of widespread deployment internationally as early as 1997; each successive year, numerous trade journals have dubbed the ensuing year as the "year of PKI."[1] Yet full full-fledged PKI rollout throughout DoD and the business community continues to lag. The commitment to PKI is not homogeneous throughout government and business entities for a variety of factors. The most daunting, yet least discussed factor is Certificate Revocation that has been described as the "5000-pound

[1] "Only Mostly Dead", Scott Berinato, [http://www.cio.com]. Accessed May 2003

elephant in the living room that everyone is trying to ignore".[2] Without a fully implemented, reliable, and viable method to implement Certificate Revocation, the PKI Trust model cannot be expected to reach maturity. This crucial element has been implemented in many different ways; however, there is no generally accepted standardized and consistent way of implementing Certificate Revocation.

1. Scope of Assumptions

In accordance with the DoD Certificate Policy (CP), every DoD organization was mandated to deploy an infrastructure having the capability to issue Class 3 PKI certificates to each member of their organization by October 2002. PKI services are becoming increasingly important in networked environments where communications and transactions occur over unsecured channels. There are a number of different types of Public Key Crypto system implementations existing in common usage. DoD subscribes to a Public Key Crypto system as described in IETF X.509 Version 3 public-key certificate. Version 3 public-key certificates were introduced to correct deficiencies associated with Version 1 and 2 definitions. Version 3 offers improvements over the previous two implementations by adding optional extensions that may be modified for use by the specific PKI implementation. Additionally, DoD subscribes to IETF Certificate Revocation List (CRL) Version 2, which includes optional extensions that provide additional security over Version 1.

C. RESEARCH OBJECTIVES

1. Primary Research Question

What are the security-relevant strengths and weaknesses of all plausible (that is; currently operational, proposed, or theoretically possible) implementations of certificate revocation management for the DoD Public Key Infrastructure?

2. Subsidiary Research Question # 1

What role does certificate revocation play in the bigger picture of PKI?

3. Subsidiary Research Question #2

What does the current DoD Certification Practices Statement (CPS) mandate regarding certificate revocation management?

---

[2] Jim Hewitt, Certco.

4. Subsidiary Research Question #3

What does the current DoD Certificate Policy (CP) mandate regarding certificate revocation management?

5. Subsidiary Research Question #4

What methods of certificate revocation are currently being used, and by whom (not restricted to DoD)?

6. Subsidiary Research Question #5

What additional methods of certificate revocation have been proposed even though they are not being used?

7. Subsidiary Research Question #6

What additional methods of certificate revocation, if any, may exist that warrant consideration, but are not currently in use or proposed for use?

8. Subsidiary Research Question #7

What attributes of certificate revocation management are security relevant, and thus candidates for use as metrics in ordering the various certificate revocation management implementations?

9. Subsidiary Research Question #8

What is the impact on interoperability within the DoD of the various security attributes of certificate revocation management?

10. Subsidiary Research Question #9

How would all plausible implementations of certificate revocation management be ordered by security strength?

11. Subsidiary Research Question #10

Of all the plausible implementations of certificate revocation management, which are the best candidates with regard to interoperability for DoD use, in order by most to least secure?

THIS PAGE INTENTIONALLY LEFT BLANK

II. PKI REFRESHER

A. TRUST RELATIONSHIPS

Use and implementation of the Public Key Infrastructure trust model creates a substantial rift in the traditional trust model. In the traditional model, the user or organization held the responsibility for verifying the identity of the user, usually defined by an identification card, a personal meeting, or a third party counsel. The entity to be trusted was personally known and identified to the trusting organization and could be tied concretely to physical credentials. The model for Public Key Cryptography offloads this trust to a third party, the Certification Authority, which provides verification by cryptographically binding and entity's identity to a unique cryptographic key – a digital certificate.[3] Each party can then present or cryptographically employ such a digital certificate which to either proves their identity or electronically "sign" digitized information. This model removes the physical barriers as well as the time barriers needed for the trust relationship to occur. However, each party must implicitly trust the third party intermediary as well as the associated framework. For this to be accomplished, it is required that the organization be able to unambiguously and correctly associate a digital certificate with the correct entity.[4]

Widespread use of PKI to provide authentication and non-repudiation has begun at fringe elements of both business and other trust relationships. The Government, most notably in the form of the Department of Defense, mandate for the PKI will undoubtedly spearhead the efforts and cause Public Key (PK) cryptography to become more widely accepted. Eventually software will exist that will allow secure communications through any electronic medium; stakeholders will inherently trust any other entity that has completed the validation process. The opportunity for misuse of digital certificates by

[3] It is worth noting at this point that the DoD PKI X.509 is not compatible with the popular Pretty Good Privacy (PGP) due to the difference in the trust model. These differences manifest themselves in the differences in between X.509 Version 3 certificates and PGP Keys (certificates). Trust decisions in the PGP model are offloaded to individuals rather than the CA.

[4] Core PKI Services: Authentication, Integrity, and Confidentiality, November 9, 1999, [www.informit.com]. Accessed Jun 2003

many actors exists[5]. The hurdle for organizations is to implement this trust relationship securely while assuaging the fears and concerns of primary stakeholders as well as actors within the bounds of the system.

Acceptance of this model requires forward thinking organizations that will accept and implement the PKI trust model. Factors that will retard the implementation of the PKI as a trust model include:

- Fear of the new model by stakeholders
- Lack of a ubiquitous infrastructure
- Liabilities, monetarily and otherwise, which an organization may be subject to in the event of a breach or failure of PKI implementation
- Lack of resolve, on many fronts, to implement PKI
- A less than fully implemented method of Certificate Revocation

Factors that will speed the implementation of PKI as a trust model include:

- The "Digital Security Act" of Oct 2001
- Widespread PKI enabled end-user applications
- Implementers and users familiar with PKI technology and applications
- Timely and reliable methods to verify certificate validity

B. SERVICES PROVIDED BY PKI

Public Key Infrastructure, or PKI, is a set of mechanisms (laws, policy, procedures, and technologies) for the use of digital credentials, to include digital signatures and document encryption, that provides confidentiality, authenticity, integrity and non-repudiation in regards to the transmission of electronic messages and data. A digital signature, provided by PKI, serves the purpose of providing authenticity (entity authentication) while simultaneously providing integrity over the signed (attached) (unnecessary) data.[6] These topics are defined as follows.

Confidentiality –Ensures that the content of information is kept from all but those authorized to have access it. In a PKI it ensures that only the sender and intended recipients are able to view the message or data transmitted by the sender.

---

5 Rethinking Public Key Infrastructures and Digital Certificates: Building in Privacy, Steven Brands, MIT Press; 1st Edition, (August 28, 2000).

6 Core PKI Services: Authentication, Integrity, and Confidentiality, November 9, 1999, [www.informit.com]. Accessed Jun 2003.

Authentication – In a PKI, authentication encompasses both entity authentication and data origin authentication, depending on which values the digital signature is computed over. The authentication allows for verification of identities in that the transmitter and receiver of a message are the same as those that are represented in the message header. "Data origin authentication implicitly provides data integrity; if a message is modified, the source has changed."[7]

Data Integrity – Addresses the possible alteration of data between sender and receiver; to insure data integrity the intended receiver must be able to detect unauthorized data manipulation through either intentional or accidental means. Manipulation may encompass insertion, deletion, or substitution.

Non-Repudiation – Through authentication and data integrity, proof is provided regarding both the integrity and origin of data; this then positively links actions related to data to a given individual or entity. This positive linking prevents the entity from denying having performed a particular action related to data.

The X.509 (4th edition) 2000 IETF provides the framework for modern PKI systems. Through the use of PKI and digital certificates, which bind a public key to an individual, device, or organization and carry the signature of a trusted Certification Authority, the above listed services may be provided.

C. PUBLIC KEY INFRASTRUCTURE COMPONENTS

The DoD classifies four different domains of information assurance; DoD Class 2, DoD Class 3, DoD Class 4, and DoD Class 5. The intent of this paper is to research DoD Class 3 PKI certificates; however, the concepts provide some overlap into Class 2 and 4. DoD Class 2 "is intended for applications handling unclassified information of low value in a minimally or Moderately Protected Environment."[8] DoD Class 3 "is intended for applications handling unclassified medium value information in Moderately Protected Environments, unclassified high value information in Highly Protected Environments, and discretionary access control of classified information in Highly Protected

[7] Handbook of Applied Cryptography, Oorschot Menezes, Vanstone; CRC Press, 1996.

[8] Ibid., p. 5.

Environments."[9] DoD Class 4 "is intended for applications handling high value unclassified information in Minimally Protected Environments." Each of these classes use some form of PKI X.509 architecture.

The components of a DoD CLASS 3 PKI System Architecture germane to certificate revocation are:

- Certificate Authority (CA) (both root and intermediate CA's)
- Certificate Directory and Backup Directory
- Key Escrow/Certificate Revocation List (CRL) Archive Server
- Registration Authority (RA) Workstation
- Local Registration Authority (LRA) Workstation

These components are briefly described below.

Certification Authority (CA) – The CA is the entity responsible for issuing and administering the digital certificates. Certification is the act of binding a subject name with a public key. The CA acts as the root agent of trust in the PKI; each member of the chain must implicitly trust the certificates generated by the CA. The actual digital certificate is issued by the CA and authenticates the user to the digital medium. A CA performs these main functions.

- Issues users with mechanism to generate private / public key
- Certifies users public keys
- Publishes users certificates
- Issues certificate revocation lists (CRLs)
- Maintains certificate archives
- Maintains encryption key escrow services as required by policy

Registration Authority (RA) –The RA functions are entirely administrative; it is responsible for physically verifying and recording the information of an individual which is passed to the CA in order to produce a digital certificate. The verification by the RA begins the certification process with a CA on behalf of the end-user. The verification process generally occurs in a face-to-face setting and entails official documents such as identification cards. The functions of the RA may be provided as part of the CA; however, they are generally offloaded into a separate component or to a trusted third

[9] Ibid., p. 6.

party. Since the functions of a CA are limited to a specific geographic area while the entities it performs services for may be geographically diverse; this offloading of the registration process promotes a more scalable framework. This generally makes the idea of centralized registration infeasible.

Certificate Publication Point or **Directory Service** – There are two main functions:

- Publication of certificates
- Publication of Certificates through various methods which may include: Certificate Revocation Lists, Certificate Revocation List Distribution Points, XML Key Management Specification (XKMS), Online Certificate Status Protocol (OCSP), and Simple PKI (SPKI)

Certificate Revocation – Certificate revocation is the process of revoking a digital certificate before its normal validity period expires. Typical occurrences for this would be:

- A user has his private key compromised or forgotten
- A user is no longer in a trust position with his former organization and the trust mechanism must be removed

Certificate Revocation encompasses many forms enumerated in subsequent chapters. The most common notion of Certificate Revocation is the CRL, which is a binary file listing that contains the following information:

- A list of revoked certificates and the reason for their revocation
- The issuer of the CRL
- When it was issued
- When the next Version of the CRL will be published

D. X.509 SECURITY THROUGH OPTIONAL FIELDS

X.509 Version 3 certificates and Version 2 CRL's are defined by IETF RFC3280, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile". What follows is a synopsis of the pertinent parts of RFC3280 for both Certificates and CRL's. This information is required for the larger discussion of CRL implementation and security. A key component of both Version 3 Certificates and Version 2 CRL's is the implementation of optional field extensions, which are used to

incorporate additional fields into the Certificate or CRL. Optional extensions allow association of additional attributes with either users or public keys, while retaining management of the certification hierarchy.

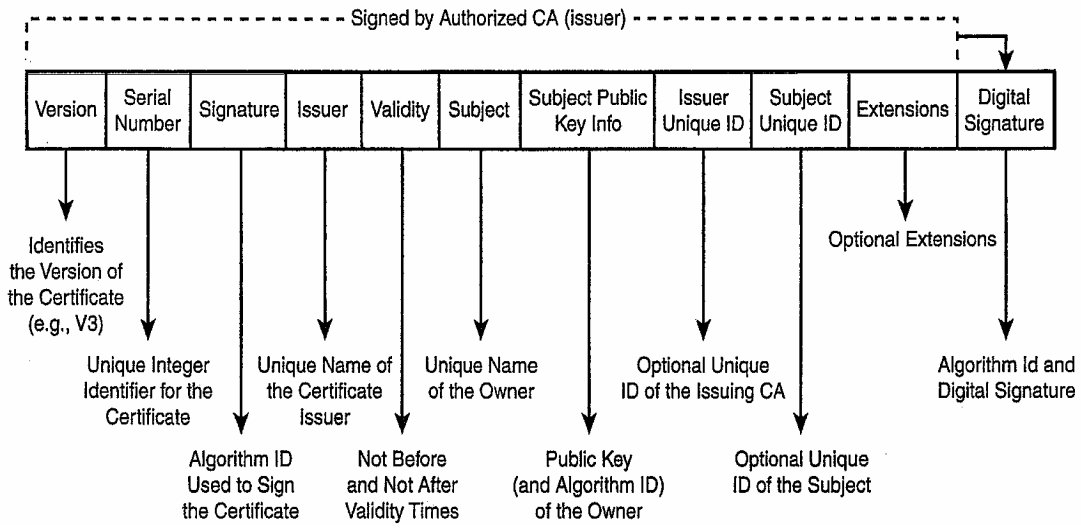

Figure 1. X.509 Version 3 Certificate.

The X.509 framework has evolved through several modifications; however, the basic model remains. Both the Certificate and CRL formats were extended through the revisions but are backward compatible through the use of the "critical" and "non-critical" designations. Several optional extensions in the X.509 Version 3 certificate are germane to the discussion of Certificate Revocation. Each optional extension contains an object identifier value that governs the basic data type (text, string, date). The optional extensions field, as seen above in the Version 3 certificate, is marked by a flag known as a Criticality Indicator. This flag indicates whether an occurrence of an extension is critical or non-critical. When an optional extension is marked critical, an application validating a certificate must process and understand the field extension. If this cannot be accomplished, the certificate must be rejected. An application validating a certificate may gracefully ignore an unrecognized non- critical CRL entry extension.

Key Usage – A critical extension that, when employed, contains bit string information which is used to define the purpose of the key (e.g. encipherment, signature, certificate signing) which is contained in the certificate. This extension MUST appear in certificates that contain public keys that are used to validate digital signatures on other public key certificates or CRL's

Authority Key Identifier – Mandated by RFC3280 for inclusion in all but self-signed certificates, this extension is a unique identifier that serves to distinguish among multiple keys supplied by the certificate issuer.

Key Usage – Sequence of one or more OID's that defines one or more purposes for the public key in the certificate (e.g. encipherment, signature, certificate signing) in addition to the usages mapped from the Key Usage field. In general, this extension will appear only in end entity certificates and is marked either critical or non-critical at the option of the certificate issuer.

CRL Distribution Point – A non-critical extension, it identifies how CRL information is obtained and the location of the CRL partition where revocation information for this certificate resides. RFC3280 provides more detailed specification for this field.

Certificate Policies – A non-critical extension it defines one or more policy object identifiers (OID's) and optional qualifiers associated with the issuance and the use of the certificate. To promote interoperability, RFC3280 specifies that the OID be used absent of any qualifiers even while the RFC defined two qualifiers. The first is the Certification Practices Statement that provides a URI at which the end user can find the CPS published by the CA. The second is the User Notice that displays information to the relying party when a certificate is used.

Basic Constraints – The basic constraints extension identifies whether the subject of the certificate is a CA and the maximum depth of valid certification paths that are included this certificate.

E. X.509 CRL STRUCTURE

The X.509 v2 CRL syntax is shown below. For signature calculation, the data that is to be signed is ASN.1 DER encoded. ASN.1 DER encoding is a tag, length, value encoding system for each element.

CertificateList ::= SEQUENCE {

tbsCertList TBSCertList,
signatureAlgorithm AlgorithmIdentifier,
signatureValue BIT STRING }

TBSCertList ::= SEQUENCE {

Version Version OPTIONAL, -- if present, MUST be v2
signature AlgorithmIdentifier,
issuer Name,
thisUpdate Time,
NextUpdate Time OPTIONAL,
revokedCertificates SEQUENCE OF SEQUENCE {

userCertificate CertificateSerialNumber,
revocationDate Time,
crlEntryExtensions Extensions OPTIONAL -- if present, MUST be v2 }

OPTIONAL, crlExtensions [0] EXPLICIT Extensions OPTIONAL -- if present, MUST be v2 }

Graphically, the CRL appears thusly:

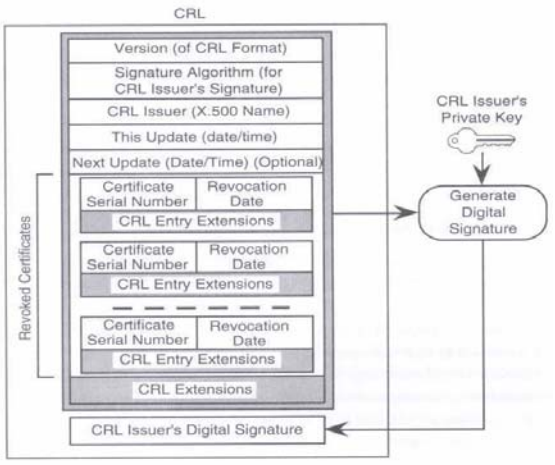

Figure 2. X.509 Version 2 CRL[10].

[10] [www.cs.odu.edu/~vbanavar/cerftificates.ppt]. Accessed Jul 2003.

X.509 Version 2 CRL fields are further described below.

Version: An optional field which describes the Version of the encoded certificate; when an extension is used it must specify Version 2 (integer value of 1) or have no value. A certificate with no value in this field indicates a Version 1 certificate that in turn indicates that this field was not defined.

Signature: An indicator of the object identifier (OID) of the algorithm identifier used to sign the CRL. PKIXALGS (RFC 3279, April 2002) provides the OIDs for the most popular signature algorithms.

Issuer: Identifies the entity (name) that has signed and issued the CRL. The issuer name field MUST contain an X.501 distinguished name (DN) and MUST follow the ASN.1 encoding structure for the issuer name field in the certificate. Implementations of this specification MUST be prepared to receive the *domainComponen*t attribute that allows for mapping of distinguished name (DN) to the Domain Name System (DNS).

This Update: Indicates the issue time and date of the given CRL; it may be encoded as *UTCTime* or *GeneralizedTime*. RFC3280 provides further instructions for the presentation of this data.

Next Update - Indicates the no later than date and time of issue of next CRL. This distinction is important as new CRL's may be issued more frequently than the specified *NextUpdate* time.

List of Revoked Certificates – The certificates on the list have the following attributes: Certificate Serial Number, Revocation date and time, Optional per-entry Extensions. Certificates are listed by serial number in ascending order. Time may be expressed in either *UTCTime* or *GeneralizedTim*e. Additional information may be supplied through the CRL per-entry extensions and are implemented on a per entry basis through per-entry extensions.

Certificate Per-Entry Extensions - CRL per-entry extensions are both separate and distinct from the CRL extensions. The per-entry extensions provide a method for associating additional attributes with specific CRL entries. The X.509 Version 2 CRL

format allows for individual definitions for a given organization or community; this is not possible with a Version 1 CRL. As with certificate extensions, each per-entry extension must be designated either critical or non-critical and follow the same compliance model for handling as certificate extensions. The recommended extensions used within Internet CRL entries and standard locations for information follow. Communities may define additional CRL extensions, however, use of critical extensions which are not standard within the larger community may cause the certificate to be rejected if read by an application which does not understand a critically marked extension. The following extension explanations are taken from RFC3280, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile".

Reason Code - A non-critical CRL entry extension, it identifies the reason for revocation of a certificate in question. RFC3280 strongly encourages CRL issuer to include meaningful reason codes. The reason code CRL entry extension SHOULD be absent instead of using the unspecified (0) *reasonCode* value. The *CAcomprimise* revocation reason applies only to attribute certificates.

```
id-ce-cRLReason OBJECT IDENTIFIER ::= { id-ce 21 }

-- reasonCode ::= { CRLReason }

CRLReason ::= ENUMERATED {

        unspecified (0),

        keyCompromise (1),

        cACompromise (2),

        affiliationChanged (3),

        superseded (4),

        cessationOfOperation (5),

        certificateHold (6),

        removeFromCRL (8),
```

privilegeWithdrawn (9),

aACompromise (10) }

Certificate Issuer – This is the name of the certificate issuer and is only required for inclusion when an Indirect CRL is issued. If this extension is included, it must be marked critical; lack of a critical flag could cause an application to incorrectly attribute CRL entries to the proper certificates. RFC3280 recommends that implementations recognize this extension. This field is defined as follows.

id-ce-certificateIssuer OBJECT IDENTIFIER ::= { id-ce 29 }

certificateIssuer ::= GeneralNames

Hold Instruction Code – This extension allows a certificate to be temporarily suspended and allows for subsequent re-instantiation or revocation. The OID indicates the action to be taken after encountering a certificate that has been placed on hold. Applications which encounter an *id-holdinstruction-callissuer* must call the certificate issuer or reject the certificate. Conforming applications that encounter an id-holdinstruction-reject MUST reject the certificate.

id-ce-holdInstructionCode OBJECT IDENTIFIER ::= { id-ce 23 }

holdInstructionCode ::= OBJECT IDENTIFIER

The following instruction codes have been defined. Conforming applications that process this extension MUST recognize the following instruction codes.

holdInstruction OBJECT IDENTIFIER ::= { iso(1) member-body(2) us(840) x9-57(10040) 2 }

id-holdinstruction-none OBJECT IDENTIFIER ::= {holdInstruction 1}
id-holdinstruction-callissuer OBJECT IDENTIFIER ::= {holdInstruction 2}
id-holdinstruction-reject OBJECT IDENTIFIER ::= {holdInstruction 3}

Invalidity Date – A non-critical CRL entry extension that provides the date on which it is known or suspected that the private key was compromised or that the certificate otherwise became invalid. The *GeneralizedTime* values included in this field must be expressed in Greenwich Mean Time (Zulu), and must be specified and interpreted as defined by RFC3280.

id-ce-invalidityDate OBJECT IDENTIFIER ::= { id-ce 24 }

invalidityDate ::= GeneralizedTime

Per-CRL Extensions - Only implemented in Version 2 certificates, these additional fields relate to the implementation of the CRL vice the certificate and provide methods for associating additional attributes with the CRL. As with many other sections of the RFC3280 specification, the Per-CRL Extensions may be added to by specific communities through the use of private extensions. The critical fields caution regarding the creation of private extensions that may not be understood in a general context remains in effect. Marking of private extensions as critical, unless critical to the workings of a specific domain, is highly discouraged.

RFC3280 does not define any private CRL extensions; the per-CRL extensions as defined by RFC3280 are as follows: Authority Key Identifier, Issuer Alternative Name, CRL Number, CRL Scope, Status Referrals, CRL Stream Identifier, Ordered Lists, Delta Information, Issuing Distribution Point, Delta CRL Indicator, Base Update, and Freshest CRL. CRL issuers are required to include the following extensions:

Authority Key Identifier – A non-critical field that identifies the unique public key that corresponds to the private key used to sign the CRL in question. It distinguishes between the multiple possible public keys that were published by the CA. The identification may be based upon either the subject key identifier in the CRL signer's certificate or on the issuer name and serial number.

CRL Number – A non-critical extension that assigns a serial number to the CRL. The number is a monotonically increasing integer for a given CRL scope and CRL issuer; it allows for detection of missing CRL's as well as determination of CRL supercession. The CRL Number must be unique to an issuer; RFC3280 suggests that issuers as well as

end applications must be implemented with the capability of handling numbers longer than 20 octets. If a CRL issuer generates two CRL's (two complete CRL's, two delta CRL's, or a complete CRL and a delta CRL) for the same scope at different times, the two CRL's must not have the same CRL number. If the "this update field" in the two CRL's are not identical, the CRL numbers must be different

Fields which are not mandated for inclusion by RFC3280 but require further explanation in the context of this document follow:

CRL Scope – Defines the method in which a CRL has been partitioned. This is a critical extension.

Delta CRL Indicator - A critical CRL extension, it identifies a CRL as a Delta CRL. The Delta CRL Indicator contains the value of the Base CRL Number as well as all update information for a Delta CRL; this information taken together defines the total information in the CRL repository.

Issuing Distribution Point – A critical extension that identifies the CRL distribution point and the types of certificates the particular CRL is designed to revoke (e.g., end entity, CA, attribute certificates)

Freshest CRL – Also known s a Delta CRL Distribution Point, this non-critical extension identifies or points to the most recent (freshest) information available. In practice, the distribution point name provides the location at which a delta CRL for this complete CRL can be found. This is not used to validate the CRL or the referenced delta CRL's; instead it is only a pointer to the most recent information.

THIS PAGE INTENTIONALLY LEFT BLANK

III. IMPLEMENTATIONS OF CERTIFICATE REVOCATION MANAGEMENT

A. CERTIFICATE REVOCATION MECHANISM DISCUSSION

CRL issuers issue CRL's; generally, this is the CA. The CRL is published to provide updated status regarding the certificates issued by the CA. A CA may delegate responsibility to publish a CRL to another trusted party. When a CRL is published by a party that is not the CA that issued the certificates, it is referred to as an Indirect CRL[11]. For qualitative and quantitative judgments to be made regarding the efficacy and efficiency of differing revocation mechanisms, it is useful to group the different mechanism. This paper uses a loose classification scheme that breaks the methods into three classes: periodic publication methods, online checking methods, and hybrids.

Periodic publication methods are by nature a static scheme in which revocation is published by a CA. The information is signed by the CA to ensure its integrity.

Online checking methods provide certificate revocation methods in a trusted directory; latency and validity checks are performed during the pulling of information from the directory.

Hybrid classes encompass elements of either of the preceding two methods but cannot be fully classified by either.

B. PERIODIC PUBLICATION METHODS

1. Certificate Revocation Lists

Complete Certificate Revocation Lists are simply digitally signed lists of all previously valid certificates that have been revoked before their scheduled expiration period. End users check the validity of presented certificates against this list. The CRL can either be "pulled" or "pushed" to the end user. In the pull model, the end-user downloads the entire CRL from the CA. The downloaded list may be cached or saved in another method location for offline verification of presented certificates. The push model

[11] IETF RFC3280, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", p. 48, Section 5.

is identical to the pull model in the information that is obtained; the difference is the method in which it is obtained. In the push model, the onus is on the CA to send the entire CRL to the end user.

Both models have in common the requirement that complete lists are published at known periods as specified by the CA. The timing of the list can be several hours to a day depending on the needs of the organization. A longer time between publishing leads to reduced confidence in the validity of the presented certificate. In circumstances as deemed by the CA, the CRL can be published more frequently than the specified period.[12] Depending on the size of the CA and the size of the organization, a complete CRL can grow to several megabytes in size. Some CA's push the complete CRL to distributed sites where they are pulled by the end-user. Similarly, some organizations pull the CRL information and store it on organizational repositories from which the end-user may use the information to verify presented certificates. The limiting factor in the basic CRL model is the sprawling growth of the CRL. A good rule of thumb for CRL size estimation is 51 bytes for the CRL structure data and 9 bytes for each revoked certificate contained on the list. A CA with thousands of end users can find that they must provide a mechanism to push or pull several Megabytes (1,048,576 bytes) of data to provide for revocation checking needs. The basic CRL provides the framework to explore revocation schemes that seek to improve on this model. It is important to note that the decision of which scheme to implement is nearly irrevocable. The chosen PKI implementation may be modified in some instances, however, once a scheme is in place it generally cannot be modified without major changes to the entire implementation.

2. CRL Distribution Points or Partitioned CRL's

Known by either name, the CRL Distribution Points allow the CRL data to be stored in a distributed format by employing one or more CRL's as well as storing the various CRL's on multiple servers. This scheme allows partitioning of the CRL into more manageable pieces. The end user certificate has an embedded pointer that points the certificate verifier to the CRL Distribution Point that is then redirected to the correct

[12] [http://www.rsasecurity.com/rsalabs/faq/4-1-3-16.html]. Accessed Aug 2003.

CRL partition based upon this information.[13] This may be leveraged further for large CA's; that is, users may be partitioned into logical functional groups with a CRL partition dedicated to each group. The CRL Distribution Point offers a more scalable method of implementing a CRL; however, the partition points are static and do not allow for changing change as the organization or CA changes.

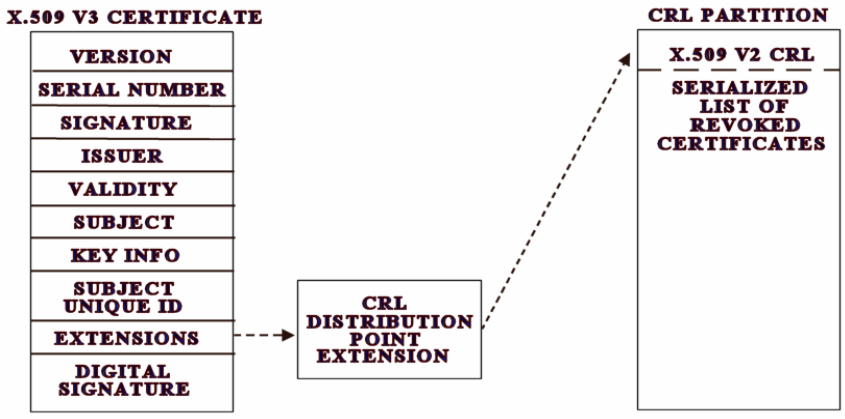

Figure 3. Redirect CRL.

3. Redirect CRL's and Enhanced CRL Distribution Points

A flaw in the CRL Distribution point is the static nature of the CRL partition. Once a certificate is issued and a CRL Distribution pointer defined, the location of the CRL Distribution Point as well as the format of the partition must remain static in order to retain the association with the certificate. This limits both the flexibility and scalability of the CA revocation scheme. Both the Enhanced CRL Distribution Point and Redirect CRL seek to rectify this situation but do so with different implementation schemes. Each scheme creates a dynamic CRL partition that allows for redirection of CRL inquiries after the certificate is issued thus retaining the CRL and Certificate association. This can be accomplished by having the originally associated CRL redirect inquiries to the appropriate partition.

The Redirect CRL uses a X.509 Version 3 Certificate extension, the CRL Distribution Point Extension, and a X.509 Version 2 CRL extension, the Redirect Pointer,

[13] [http://www.verisign.com.au/whitepapers/enterprise/revocation/cert_revk2.shtml]. Accessed Jul 2003.

to provide the location of the partition associated with the certificate revocation information. The Certificate CRL Distribution Point Extension points to a Redirect CRL which contains a X.509 Version 2 CRL with a valid Status Referral extension. The Status Referral Extension in turn points to the correct, and possibly dynamic, CRL partition.

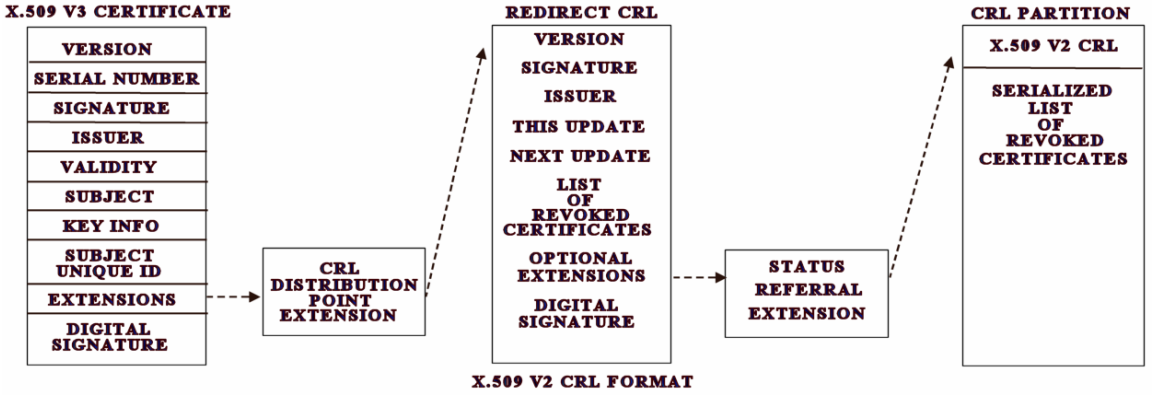

Figure 4. Redirect CRL.

The Enhanced Distribution Point CRL, suggested by the IETF PKIX Working Group in 1998, separates the location and validation functions by using the CRL extensions Status Referral*s* and CRL Scope.[14] The Status Referrals extension is used to convey the location of the location of the latest CRL. After the end user application has successfully located the CRL, the CRL Scope is used to determine whether the located CRL's contain information appropriate to the status of the end user certificate.

4. Delta Certificate Revocation Lists (CRL)

The Delta CRL is composed of a base list and updates. The base list is a complete CRL as of a defined time period. The update or Delta CRL contains incremental CRL information; Delta CRL's may be formatted relative to a base CRL or relative to a particular point in time.[15] The "Delta CRL Indicator" extension is used to denote which method is used. The Delta CRL allows the end user who has retrieved the latest full CRL to retrieve a smaller amount of updated CRL information thus maintaining a complete list

[14] Enhanced CRL Distribution Options, IETF PKIX Working Group, August 7, 1998, [http://www.ietf.org/internet - drafts/draft-ietf-pkix-ocdp-01.txt]. Accessed Jul 2003.

[15] Understanding PKI, p. 118.

of revoked CRL's and increasing the confidence in the PKI. The Delta CRL allows for more frequent publishing in an attempt to optimize the timeliness of available information against the required bandwidth for transmission or retrieval by the end user.

5. Indirect Delta CRL's

The Indirect Delta CRL enables multiple CA's to maintain information on one CRL. It is similar to the Delta CRL in that it has a base, however, the CRL information is from one or more CA's and the information contained is signed by several entities. This method allows for simplification of retrieving certificates in the cross trust model.

6. Sliding Window Delta CRL

A problem with Delta CRL's is the rate of request for full CRL's in relation to the rate of request for Delta CRL's. With a traditional Delta CRL, the expiration date for the retrieved CRL is governed by the NextUpdate field in the CRL. The end user will request Delta CRL's until the NextUpdate time in the cached CRL is reached. After the newest full CRL is published, the users seeking to verify certificates will seek to obtain the full CRL, and statistically the majority of users in a system will seek a full CRL within a relatively short period centered around the arrival of the *NextUpdate* publish time. This period of time represen*t*s a spike in system and resource utilization which can adversely affect the timeliness of revocation checking should be minimized. The graph in Figure 5 demonstrates this concept (This graph indicates the probability of request for a full CRL for 30,000 end users requesting 10 certificates per day with a full CRL issued only at time zero).

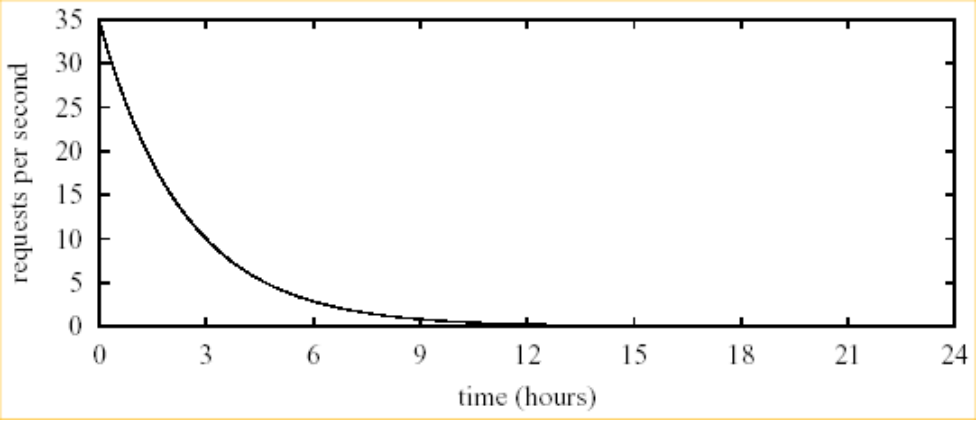

Figure 5. Un-Segmented CRL[16].

[16] "A More Efficient Use of Delta CRL's", David Cooper, p. 2, [http://csrc.nist.gov/pki/documents/sliding_window.pdf]. Accessed Mar 2003.

Reducing the ratio of full CRL requests to Delta CRL requests reduces the instantaneous load on the repository and improves response time. A simplified method to achieve this is over-issuing of CRL's; the CA issues new CRL's more often than necessary. In an over-issued CRL the CA might issue a new CRL four times a day when the CRL has a life span of 24 hours. In this method the end users will have full CRL's that expire at different times; over-issuing CRL's causes the distribution of requests for full CRL's to be more evenly distributed and lowers the peak request rate, as can be seen in Figure 6.

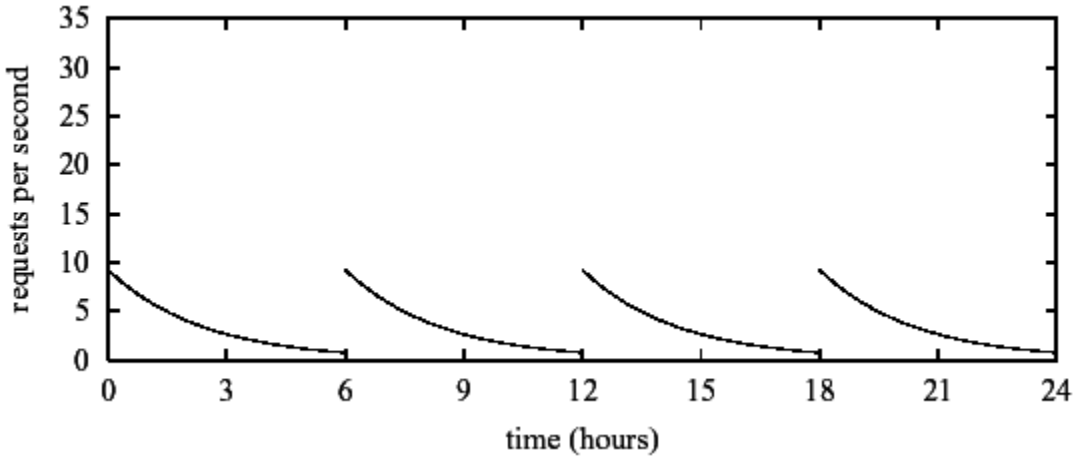

Figure 6. Over-Issued CRL[17].

The *BaseCRLNumber* provides information to the end user application as to whether a certificates status has changed between the time the full CRL is issued and the issuance of a Delta CRL. The historic use of Delta CRL's has been to issue CRL's based on the numbers of certificates expired, vice rather than on the passing of a given specific time period. Sliding Window Delta (SWD) CRL's, as outlined by David Cooper of NIST, is an improvement on both the Delta CRL and over-issued Delta CRL. The SWD CRL reduces the rate of request of for full CRL's by implementing fixed windows that are sufficiently large enough to provide up to date CRL information without causing repeat requests for base CRL information.[18]

[17] "A More Efficient Use of Delta CRL's", David Cooper, [http://csrc.nist.gov/pki/documents/sliding_window.pdf], p. 3. Accessed Mar 2003.

[18] Ibid.

7. Certification Authority Revocation Lists (CARLs) or Authority Revocation Lists (ARLs)

Known by either name, CARLs are used exclusively to publish the serial numbers of CA public key certificates, including cross-certificates, which have been revoked. The function of these certificates is identical to that of the end user certificate. However, the intended user is different. The CARLs may be implemented in any of the schemes appropriate for an end user X.509 V2 CRL certificate. This type of mechanism is required due to the cross certificate / (i.e., "cross trust") nature of PKI. It should be very rare that a CARL is issued.

8. End-Entity Public-Key Certificate Revocation Lists (EPRLs)

The opposite of a CARL, the EPRL contains certificates issued only to individuals; CA revocation information is not included in this list.

9. Certificate Revocation Trees (CRTs)

The CRT is a proprietary ValiCert technology that was introduced by Paul Kocher in his 1988 paper "A Quick Introduction to Certificate Revocation Trees" and is a departure from the CRL structure discussed thus far. Valicert developed and implemented CRT's but stopped supporting CRT's approximatley three years ago.[19] The CRT is designed to provide the end user with a short proof that the certificate in question had not been revoked while providing revocation information for one or more CA's or communities. CRT's use a Merkle hash tree in which a binary tree represents all known certificate revocation information for the specified CA's or community.

> The issuer sorts the list, optionally removes any duplicate entries, then adds a beginning-of-list marker and an end-of-list marker. Each pair of adjacent entries in this sorted list specifies a range between which there are no list entries. Except for the beginning and end markers, each list entry appears in two ranges, once as a minimum value and once as a maximum value. A hash tree is then constructed where leaf nodes correspond to ranges in the list. Because the tree's leaf nodes define intervals, this structure is referred to as an interval hash tree.[20]

The hash tree is formed by hashing each set of statements; this then forms the leaves on the tree. The combination of each node of the next level (descendant) is

---

[19] John Hines, Engineering Director, Valicert/Tumbleweed Communications.

[20] US Patent 642689, US Patent Office.

computed by hashing the concatenation of its ancestor.[21] The resulting set of statements are hashed to form branches, this process is carried out until the Root Node is formed by combining and hashing adjacent branches which then forms a lower level. The Merkle hash tree is hash tree is used because it reduces the cost per signature over a large number of signatures by combining a large number of items into a single root node that can be digitally signed. This method provides assurance that the all items contained within the tree have been signed.

A graphic depiction of the CRT is in Figure 7. $N_{0,X}$ indicates a particular node or leaf of the tree that corresponds to a statement regarding an upper and lower limit of revoked certificates, where 0 is the current level of the tree and X is the given statement. The process of concatenation and subsequent hashing proceeds from left to right until the Root node is computed. The root node is combined with issuing and expiration information and digitally signed by the CRT originator.

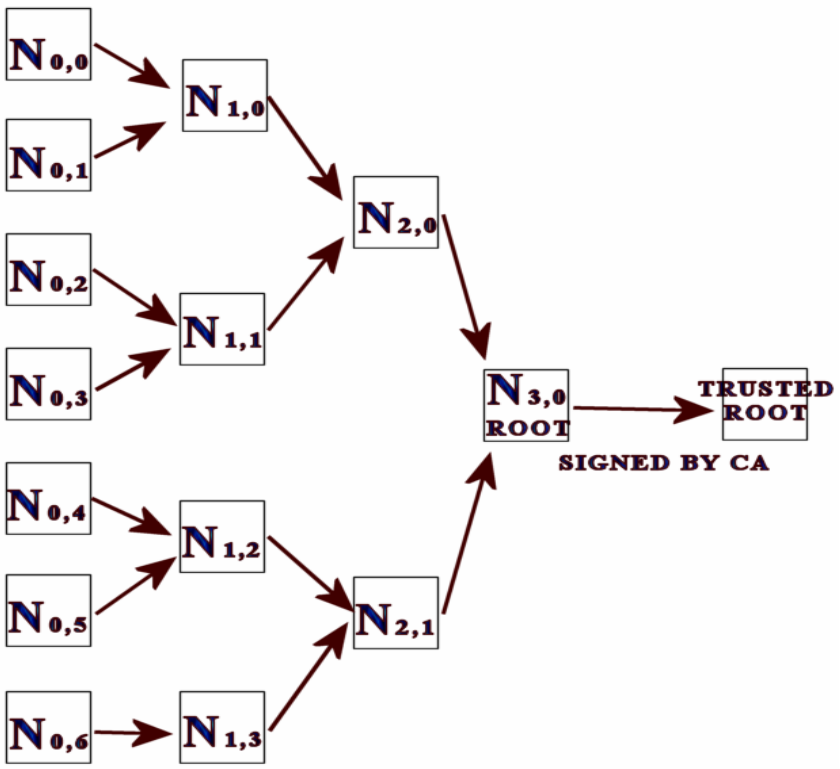

Figure 7. CRT.

[21] Ibid.

To determine if a certificate is revoked, the certificate in question is requested from the CA or repository. When responding to a verification query, the CA or repository determines the relevant leaf and returns information that will allow the relying party to determine if the leaf is within the bounds of one a statements. The following information is provided: the CA's signature on the root node, the relevant leaf, the relevant leaf's sibling, and the sibling of each of the relevant leaf's ancestors.

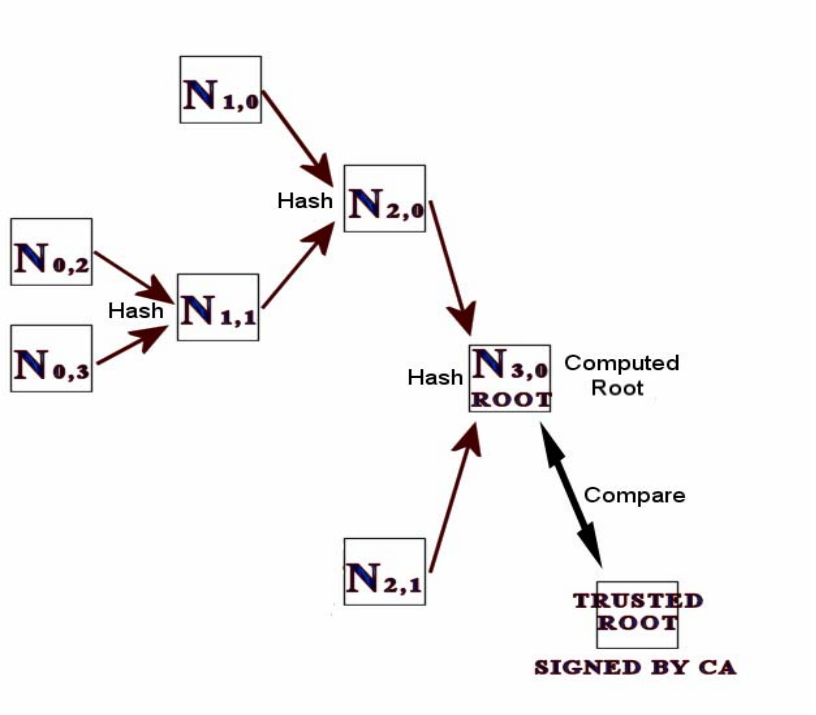

Figure 8. CRT Validation Process

As depicted in Figure 8, the relying party or a mechanism on their behalf takes information supplied by the CA or repository and hashes the leaf in question along with it sibling; this result is then the supposed parental hash that is hashed with the antecedents through the chain until a supposed root node is returned. The root node computed by the relying party mechanism is compared against the root node provided in the initial request. If these values are the same, it can be assumed that the original leaf was a part of the CA tree. This process may be completed by the end user or off-loaded to a trusted third party. .

10. MiniCRLs

Developed by Corestreet Ltd, the MiniCRL seeks to provide "ultra-low bandwidth certificate validation" using three percent of the bandwidth required for a full X.509 CRL.[22] This is not a X.509 v2 use of certificates but a new proposed standard. The MiniCRL works by pairing down the information not necessary for revocation; discarding everything except: revocation date, reason code, and invalidity date. This resultant information is then further segmented to further reduce the size of the total CRL. It is estimated that the segmentation achieves 6 bits per revoked certificate vice the 22 bytes of a traditional CRL.[23] The MiniCRL segmentation is accomplished by creating a header with administrative data and adding segments with information on a sequence of revoked certificates. An example of a proposed MiniCRL appears in Figure 9.

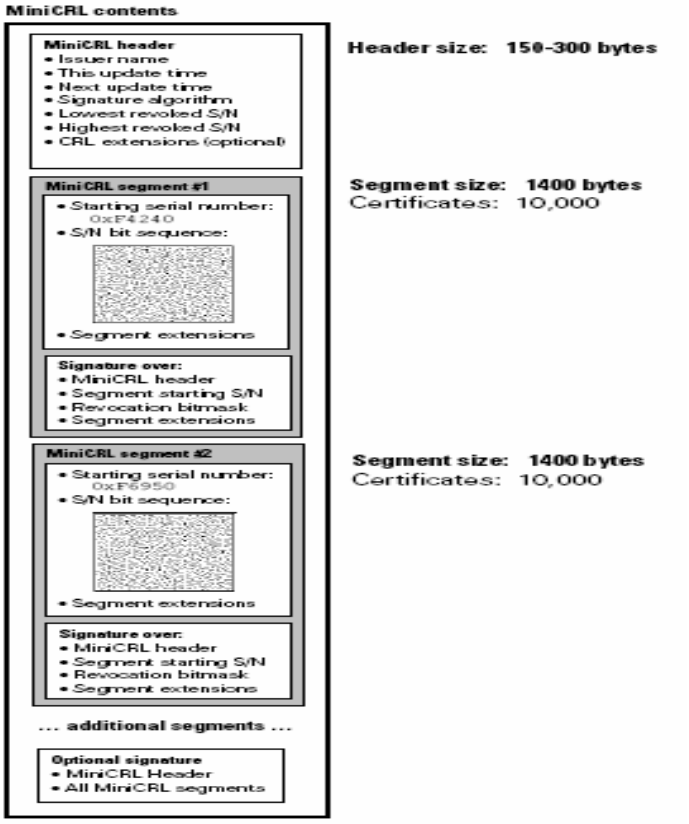

Figure 9. MiniCRL Structure[24]

[22] Corestreet. MiniCRLs: Ultra Low Bandwidth Certificate Validation. 2003. [www.corestreet.com]. Accessed Aug 2003.

[23] Ibid.

[24] Ibid.

Each segment includes a starting serial number and a serial number bit sequence. This bit sequence is a special encoding mechanism to indicate the serial number of the first certificate in the sequence, and then uses one bit to represent each serial number increasing from that starting point. Segments are added until the universe of know sequential revoked certificates are exhausted. The entire MiniCRL is then signed by the CA. For an implementation having a revocation rate of 10%, the size reduction is approximately 18:1; the MiniCRL then proposes using standard compression technology to achieve a size reduction of approximately 30:1.[25]

11. Trusted Directories

These are applicable only for intranet applications such as managed company intranets. By having all certificates available on an enterprise directory the end user checks for the availability of the certificate in question in a central directory. Revocation is as simple as deletion of a certificate by the IT manager. End applications may be designed to check for the presence of certificates in the directory prior to relying on them; this removes the requirement for revocation checking.

C. ONLINE QUERY MECHANISMS

1. Online Certificate Status Protocol (OCSP)

OCSP was defined by IETF RFC 2560, "X.509 Internet Public Key Infrastructure Online Certificate Status Protocol – OCSP." This RFC outlined Version 1 of the OCSP protocol; subsequently an internet draft protocol which outlined Version 2 was released. The latest release of IETF OSCP Version 2 was released in January 2003, it has not been finalized and while it conforms to RFC 2560 it has not yet superceded it and remains a "work in progress".

OCSP is a request-response protocol that facilitates real time communication between a client application that requests certificate revocation status and a responder server that communicates with any number of back-end revocation strategies; the back-end may be a mix of other OCSP responders, CRL's , or legacy databases. The OCSP server (or OCSP responder) is a trusted entity that provides online revocation status and may be configured to provide additional status information beyond that available through a CRL. The information that the OCSP responder uses to provide a response may be

[25] Ibid.

from several different methods: OCSP responders are configured to harvest information. Possible methods include: the responder may be configured to download CRL's and Delta-CRL information from the CA, the CRL information may be pushed to the OCSP responder by the CA, or the OCSP responder may have a database connection to the CA list of revoked certificates. The use of OCSP may be used as either a replacement of, or a supplement to checking against a CRL.

An OCSP session is straightforward; the end user application requests certificate status for a given certificate from an OCSP server. The OSCP request consists of the protocol Version, the service request type, one or more certificate identifiers, and any optional extensions that may be processed by the OCSP Responder. "The certificate identifier consists of the hash of the certificate issuer's Domain Name, the hash of the issuer's public key, and the certificate serial number."[26] Upon receipt of the request, the OCSP responder server makes a determination to provide service based upon three conditions: a well-formed request, if whether the server is configured to provide the requested service, and whether the server contains the information needed to fulfill the request for certificate status. If any of these conditions cannot be met, the OCSP responder responds with an error message. If the preceding conditions are met, the OSCSP responder responds with a digitally signed response indicating that the certificate in question is "good", "revoked" or "unknown". A "good" status indicates that, at a minimum, the certificate was not revoked at the time of the request. It does not indicate whether the CA issued the certificate, nor does it check against the validity time period. Optional extensions may be used to indicate additional information that may include issuance, validity period, etc. (validity period checking can, and should, be done by the relying party… it is simply a matter of looking in the certificate.) A "revoked" status indicates that the certificate has been revoked, either permanently or temporarily, and provides the time that the revocation occurred. Optional extensions may indicate the reason for revocation. An "unknown" status indicates that the responder does not have information about the certificate in question.

[26] Understanding PKI, p. 123.

Infrastructure implementation of the OCSP responder may be achieved several ways depending on the desired goals of the organization. Three implementations are: Self-Signed Trust, Delegated Trust with Root OCSP signing certificates, and Direct Trust OCSP Responders. A graphic representing these different implementations in the DoD framework appears in Figure 10.

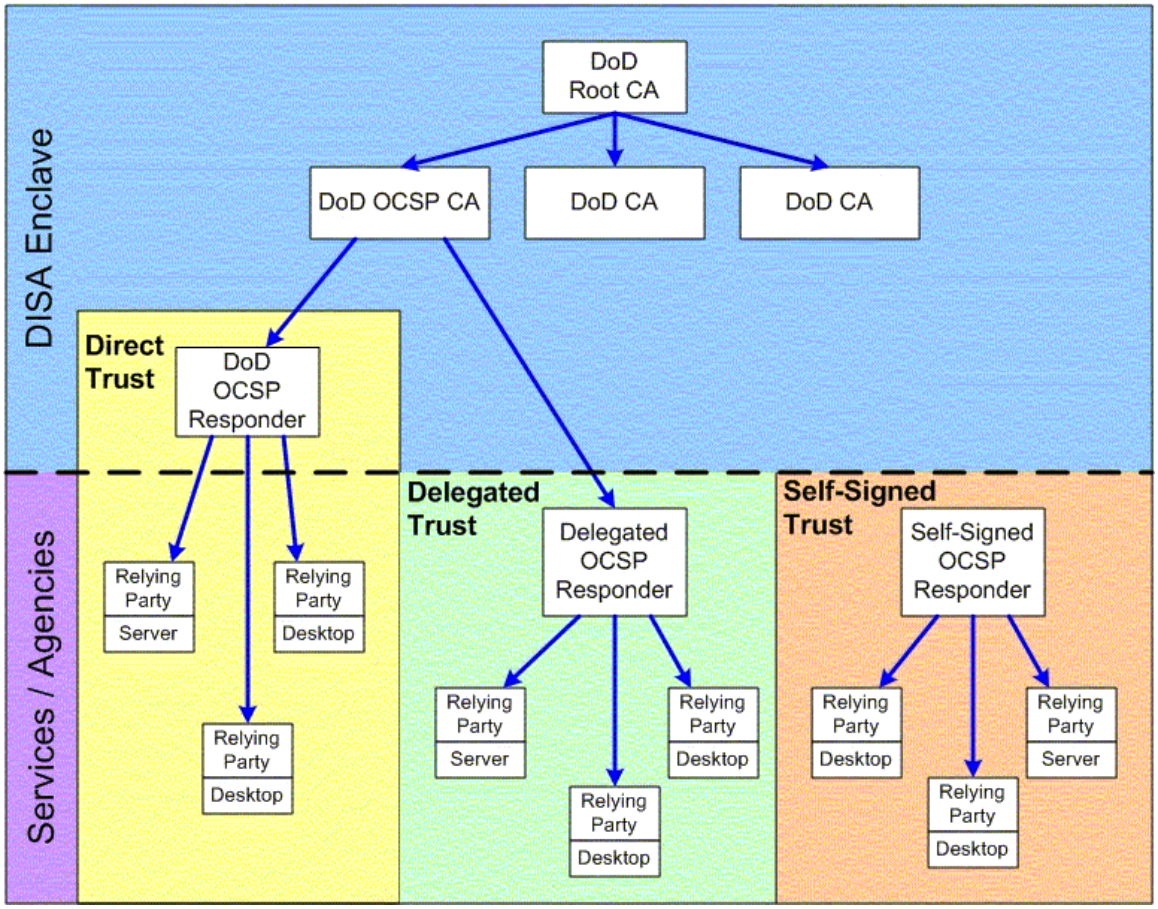

Figure 10. OCSP Implementations in DoD Framework[27].

The Self-Signed Trust implementation consists of an OCSP responder outside of the CA enclave. Certificate revocation information is harvested by either push or pull to the responder. The responder signs all requests for information with a signature from the responder; there is no signature from the root CA indicating that this server is acting on behalf of the CA organization. End user applications are configured to make requests to one or more OCSP responders as specified by the end-entity organization. Both the

[27] DMDC DoD PKI Business Working Group.

Delegated Trust and Direct Trust implementations require that the CA have a DoD OCSP CA. In the Delegated Trust, the OCSP server harvests information from the DoD OCSP CA but also receives an OCSP certificate that allows chaining of trust from the requesting entity through to the CA. End user applications are configured to make requests to one or more OCSP responders as specified by the end-entity organization. The Direct Trust model places the OCSP responder within the CA enclave. The end user applications are configured to make requests to one or more OCSP responders within the CA enclave; certificate traceability is direct to the CA.

2. Simple Certificate Validation Protocol (SCVP)

The IETF PKIX working group is developing SCVP; the latest RFC Internet-Draft memo is dated October 2003[28]. The abstract for SCVP is as follows: "SCVP allows a client to offload certificate handling to a server. The server can provide the client with a variety of valuable information about the certificate, such as whether the certificate is valid, a certification path to a trust anchor, and revocation status. SCVP has many purposes, including simplifying client implementations and allowing companies to centralize trust and policy management."[29] SCVP implementation is progressing through several end providers. The protocol uses a request and response model similar to OCSP that uses two request-response pairs. The primary request-response is used for certificate validation while the second is used for validation policy determination. SCVP is intended to reduce the amount of processing end-user applications must perform and is particularly targeted for cell phones, mobile wireless devices, or other similarly bandwidth and processor limited applications and technologies. SCVP seeks to reduce the overhead of two classes of end user applications. For the first class of application, SCVP can provide information as to whether the certificate is the proper type for the intended usage and build and check the validation path to ensure the public key belongs to the identity named in the certificate. The client delegates this to the SCVP server. The second possible application is class of application seeks certification path validation when the application has no method of constructing the validation path. The construction of this path is offloaded to the SCVP server.

---

[28] [http://www.ietf.org/internet-drafts/draft-ietf-pkix-scvp-13.txt.] Accessed Jul 2003.

[29] Ibid.

"The primary goals of SCVP are to make it easier to deploy PKI-enabled applications and to allow central administration of PKI policies within an organization."[30] In relation to certificate revocation, the SCVP servers may be implemented as a trusted or untrusted responder and may support the following types of revocation methods:

- Full CRL's
- OCSP responses
- Delta CRL's and the relevant associated full CRL's
- Any available revocation information has to be collected
- Instance in which no revocation information need be collected

It should be noted that SCVP is not so much a new protocol for validation but is more an intermediary protocol as seen in Figure 11.

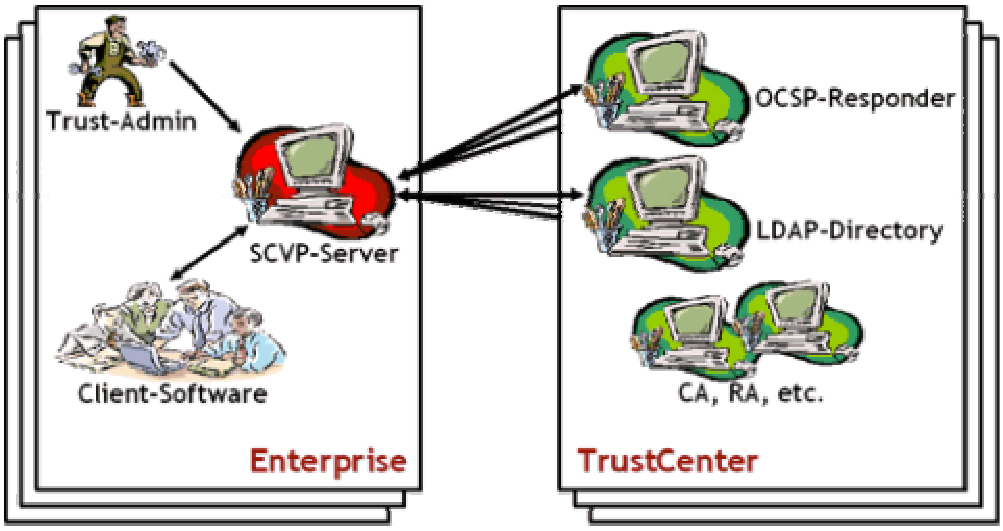

Figure 11. SCVP as an Intermediary[31].

Future SCVP implementations might include a framework that incorporates OCSP certificate revocation responses in conjunction with trusted and untrusted SCVP servers while providing information regarding certification paths, certification chains, certification path construction, and other revocation status information. SCVP is still in the IETF draft stage and has no applications built to its specifications at this time.

[30] Ibid.

[31] [www.openvalidation.org]. Accessed Jun 2003.

3. Novomodo

Developed by Silvio Micali and presented at the 1997 RSA Security Proceedings, Novomodo is a X.509 version 3 compliant small bandwidth certificate validation scheme. Novomodo is based upon mathematical one way hashing function; the CA releases a proof, either daily or at any specified time, which the relying party may use to mathematically test to determine whether the certificate in question is still valid. As part of the mechanism for issuing a certificate, the CA randomly chooses two different 20-byte values, Y0 and X0. From these, original values, hashes are obtained. Y0 is hashed once while Xz where z is the numerical number of days in the life of the certificate. These two values become validity targets where Y1 is the revocation target and X365 validity target (where X365 is equal to a life of 365 days in a scheme where the CA performs notifications once daily). The values of Y1 and Xz are included in a certificate while the 20 byte intermediate values (Y0, X1…X(z-1) are kept secret.

The CA issues a proof, either daily or as specified by the CP and CPS, which contains the information necessary to determine the status of the certificate. For day i after the issuance of the certificate, the CA releases a proof. If the certificate has been revoked, the CA releases Y0. If the certificate remains valid the CA releases X365-i. Y0 is the H-inverse of the revocation target Y1 and X365-i is the H-inverse of the validity target X365. This data may be disseminated through an online validation request method or may be posted through a periodic publication method. The relying party will be in receipt of one or two values; Y1 or Xz-i. They relying party verifies either the validity target or revocation target by either hashing Y0 once and comparing that value to the revocation target Y1 or hashing Xz-i i-times and comparing that to Xz. By hashing the data and comparing the values to either the validation target or revocation target the relying party can trust the veracity of the proof they have received.

D. HYBRID MECHANISMS

1. Micro-CRL (Flanigan Method)

During a break at the 2nd Annual PKI convention at NIST, Professor William Flanigan described a method he termed the “Micro-CRL”. To date this has not appeared in any IETF working group nor does it appear to be formally documented. In the Micro-CRL method, at the time of signing the sender pulls the most recent revocation

information regarding his certificate from the responsible CA. If the certificate has not been revoked, the CA provides the serial numbers of the closest upper and lower adjacent revoked certificates, adds a nonce to prove that the information was valid as of a given time, and signs the information with the CA private key. This certificate status information is hashed with the message that has been signed. When the receiver receives the certificate, he decrypts the message using the sender's public key and then checks revocation status using the CA public key. This method puts the onus of certificate revocation on the sender and reduces the overall bandwidth required by the receiver to process the revocation transaction.

2. Short Lived OID Certificates

In certain circumstances, the X.509 v2 certificate attributes may be set such that the life of the certificate is a relatively short length of time with the assumption that the possibility of key compromise before the expiration of the validity period is unlikely. This method could best be used in a closed environment where certificates are validity period is short lived and used as a time stamp; both the sender and the receiver are able to quickly determine whether the certificate is valid based upon the time the certificate is received. VeriSign has implemented a form of this method with its VeriSign Short-Lived Wireless Server Certificates; these certificates enable digital certificate validation for the mobile wireless Internet devices. This is defined by the WAP Public Key Infrastructure (WPKI) specification; there is no reason that this type of implementation cannot be implemented under the PKIX specification.

3. Defense Messaging System (DMS) Method[32]

The DMS system was comprised of an active directory and Local Authority Workstation (LAW) for Certificate maintenance functions. The DMS system employed a pull method for routine certificate revocation notification. In the event of a large or number of certificate revocations or a serious security incident, the system could be quickly converted to a push method and the Compromised Key List (CKL) was broadcast via secure message with a single CKL for the entire system. No further documentation of this system was found and is believed to have been decommissioned.

---

[32] [http://www.chips.navy.mil/archives/94_oct/file1.htm]. Accessed Jun 2003.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. CERTIFICATE REVOCATION MANAGEMENT POLICIES

A. CERTIFICATE PRACTICES (CP) AND CERTIFICATION PRACTICES STATEMENT (CPS)

The Certificate Policy (CP) and Certification Practices Statement (CPS) are the two primary documents that define a set of rules and agreements by which the CA and the end user agree will abide. RFC 2527 "Certificate Policy and Certification Practices Framework" is the defining source for these two documents; each document is a high level description of responsibilities and assumptions which form the relationship between the CA and associated entities and the organization implementing a PKI. The PKI community is moving toward generating a standard CPS and CP for given infrastructures; this end goal has not yet been reached and organizations and CA's must draft documents that complement each other. To ensure that the requirements outlined in the documents remains in the realm of plausibility there is usually significant interaction between the community of end users and CA's.

X.509 defines a CP as "a named set of rules that indicates the applicability of a certificate to a particular community and/or class of application with common security requirements." The CP includes statements of the certificate usage as well as the binding set of rules for certificate holders. The CP is written by the organization that will be relying on the certificate and services provided by the CA and associated entities, although the CA may write a sample CP that matches its CPS. The CPS is a statement of the practices that the CA and its associated entities will follow in deploying and managing certificates. The CPS states how it will adhere to the policies delineated by the CP. These two documents form the basis of trust between otherwise unrelated entities; both must ensure completeness as well as a full understanding and agreement of the operating practices of both parties. Lack of this understanding can cause inadvertent breaches of trust when each party believes it is living up to the standards set forth in its document.

RFC 2527 originally sought to answer any questions regarding the PKI Implementation through the detailed technical and legal aspects of certificate issuance through the Certificate Policies (CP) and Certificate Practice Statements (CPS). As PK

technology has matured, organizations implementing PKI found two problems with these documents. First, many PKI users, especially consumers, found the documents too difficult to understand. Second, the information contained in the documents provided in-depth information regarding operating procedures to parties that meant them ill. The PKI Disclosure Statement (PDS), which began as a separate IETF working document and was subsequently absorbed into RFC 2527, fills both of these voids by "disclosing critical information about the policies and practices of a CA/PKI."[33] The PDS contains and emphasizes information covered in detail by the associated CP and/or CPS documents. RFC2527 specifically indicates the PDS is designed to "act as a summary of information about the overall nature of the PKI…its purpose is to distill information about the PKI, as opposed to protecting security sensitive information contained in an unpublished CPS, although a PDS could also serve that function."[34] While RFC does not specifically indicate that the PDS cannot be used as a security device, this has become a defacto practice. While a CP is generally eight to ten pages and a CPS may approach forty or more pages, the PDS is generally one or two pages and may contain links that refer the end user to sources to retrieve greater detail if needed.

B. DOD CPS

This information is contained in the document, Defense Information Infrastructure Certification Authority, Certification Practices Statement for Release 3, Version 4.1, dated May 15, 2002. The DoD Policy Management Authority (PMA) is responsible for the review of the aforementioned CPS as well as analyzing the CPS of any commercial Certification Authorities that offer services DoD to ensure that compliance with DoD Certificate Policies. In keeping with the standard model of PKI, the DoD CA's are charged with issuance and management of certificates.

The CPS specifically states that a Certificate Management Infrastructure (CM) will be implemented Certificate and Certificate Revocation List (CRL) Generation and Distribution methods will be in place. Specific points germane to certificate revocation include:

---

[33] IETF RFC 2527.

[34] Ibid.

- The CPS is made available online at http://iase.disa.mil. This is a secure HTTP site which also requires a PKI certificate to enter the site.
- The frequency of certificate publication is defined as "specified in section 4.2.3," this section does not exist.
- Regarding frequency for publishing the CRL, "CRLs will be issued daily by the CA and every 28 days by the root CA, even if there are no changes or updates to be made, to ensure timeliness of information. CA servers automatically overwrite superceded CRL's upon posting of the latest CRL."
- Repositories are required and will consist of "a Lightweight Directory Access Protocol (LDAP) Directory Service for the purpose of publishing certificates and CRL's and a web site to publish (the) CPS."
- Revocation request processing is described in the RA CPS, LRA CPS and the VO CPS.
- There is no suspension mechanism in the DoD system. Certificates are irrevocably revoked only. CRL's will be issued every twenty-eight days by the CA. "If the CRL is issued as a (CA) result of key compromise or a CA revocation, the CRL will be posted as quickly as feasible but shall be posted within twenty-four hours after notification of the compromise or decision to revoke the CA."
- In regard reliance of the certificate for validity, "parties relying on CRL's must, via their applications, check a current, valid CRL for every certificate in the certificate path, if it is feasible to obtain these CRL's."
- The class 3 PKI currently does not support OCSP or any other form of revocation outside of CRL issuance.
- DoD policy mandates the use of X.509 Version 3 certificates and either Version 2 or 1 CRL's.

C. DOD CERTIFICATE POLICY

The DoD Certificate Policy is outlined in "X.509 Certificate Policy for the United States Department of Defense." The overview states that "DoD PKI must support five primary security services: access control, confidentiality, integrity, authentication and technical non-repudiation."[35] The CP also notes that "DOD studies have concluded that a great majority of past compromises have involved inside threats."[36] Without a working certificate revocation mechanism in place, confidentiality, integrity, authentication and technical non-repudiation cannot be adequately provided for, nor can the insider threat be

[35] X.509 Certificate Policy for the United States Department of Defense, p. 3.

[36] Ibid., p. 5.

countered. The DoD CPS asserts "The longer and more often a key is used, the more susceptible it is to loss or discovery." While the validity of this statement may be argued, the statement bolsters the argument for a fully implemented certificate revocation mechanism.

The DoD CP provides information for four different domains of information assurance; DoD Class 2, DoD Class 3, DoD Class 4, and DoD Class 5. The intent of this paper is to research DoD Class 2 and 4, although there is overlap into DoD Class 4. The policy guidance in regards to certificates is for DoD X.509 Version 3 certificates. DoD Class 3 and Class 2 infrastructures must subscribe to the Federal PKI Version 1 Technical Specifications: Part E – X.509 and CRL Extensions Profile [FPKI-E]. Both of these specifications are the Federal mandate for X.509 Version 3 certificates and Version 2 CRL's .

It is important to note that the DoD CP is not written for a specific CA; instead it is intended to map to the CPS of any DoD operated CA or the CPS of any CA that provides services to DoD entities. Since this CP is written to map to a broad range of CA's CPS, it would not be surprising to see more acceptable forms of revocation in the CP as compared to the DoD CPS. This is in fact the case; the DoD CP allows for both Certificate Revocation Lists; On-line status checking or any other method that is "described in the CA's approved CPS" as well as "providing authentication and integrity services commensurate with the assurance level of the certificate being verified." The DoD CPS allows only Certificate Revocation Lists. At a minimum, any CA which supports DoD must support CRL's, however; client software which uses online-line revocation methods is not required to additionally support CRL usage. Issuing frequency of CRL's in the CP is determined by the class of data; Class 2 allows the issuing frequency to be within a "reasonable period" as defined in the CPS of the relevant CA, Class 3 and 4 are required to issue CRL's daily. For Class 3 DoD PKI this is a direct mapping, as the CPS states that "CRLs will be issued daily by the CA and every 28 days by the root CA, even if there are no changes or updates to be made." Suspension of certificates is forbidden by both the DoD CP and CPS.

The CP states that "CA's shall make public a description of how to obtain revocation information for the certificates they publish, and an explanation of the consequences of using dated revocation information." While the methods for obtaining Revocation information are published in the DoD CPS, the consequences for using dated revocation are not. Following this line of thought, a CA supporting the DoD should have real time application notification when certification information cannot be found or is invalid. Trivial testing of Microsoft Exchange Server indicates that application notification regarding certificates is insufficient for any user to make a valid decision regarding whether to trust or deny a certificate.

D. ENTRUST CPS

Entrust Technologies provides both Secure Socket Layer Web certificates as well as PKI certificates. Very little information in the public domain exists for Entrust PKI certificates. A source within Entrust said "Entrust supports CRL's, CRL Distribution Points, XKMS, and OCSP (through partnerships)." This same source would not give any more specific information.

Information on the SSL Web certificates is available in the "SSL Web Server Certification Practice Statement"[37]. While not specifically designed for PKI, it does illustrate the technology in use at Entrust. Specifically, CRL's are implemented and updated with "reasonable efforts to issue CRL's at least once every twenty-four hours." Delta CRL's are not used, however, instances of serious compromise cause more frequent updates of the CRL. No mention of any other technology is made in this document.

E. BALTIMORE TECHNOLOGIES CPS

The information in this section is exclusively from the Boston CPS Dated 14 March 2003[38].

- Baltimore Technology Repository obligations include: Server Certificate CA, CRL distribution points for CA's, and X.500 Directory for Managed Services. CA's operating under the Boston PKI post Certificates and CRL's to them as appropriate.

---

37 [http://www.entrust.net/CPS/webcps010103.pdf]. Accessed Jun 2003.

38 [http://www.baltimore.com/omniroot/Boston_Certificate_Practice_Statement_[Company]_v1.0.pdf]. Accessed Aug 2003.

- The procedure for a suspension request, and the entities allowed to initiate a revocation request are detailed in the applicable CP or the Customer CPS.
- The CRL is updated at the CRL issuance frequency stated in the applicable CP.
- CRL checking requirements mandate that relying parties must check the validity and currency of a Certificate prior to reliance on such certificate.
- Boston operated CA's use only the X.500 Directory for CRL's.
- Boston supports and uses X.509 Version 3 Certificates which contain v.3 in the Version field and uses X.509 Version 3 Certificate Extensions.
- Boston supports and uses X.509 Version 2 CRL's and CRL Entry Extensions.
- Each CP used under the Boston PKI has been allocated an OID which provides a unique identification number for the CP and includes a policy Version number.

F. VERISIGN CPS

The information in this section is exclusively from the Verisign CPS statement of June 11, 2002[39].

- Verisign supports Certificate Revocation Lists, Partitioned CRL's, Online Certificate Status Protocol, and Trusted Directories[40], and "other value added revocation and status services".
- Verisign maintains separate CRL's for three classes of customers. For Verisign PCA's and Class 1-3 Certification Authorities, CRL's are posted in at http://crl.verisign.com. Managed PKI Lite Customer CA's have the CRL data posted at http://onsitecrl.verisign.com/OnSitePublic/. Managed PKI customer CA's have CRL information posted in customer-specific repositories.
- Verisign provides online status checking of individual certificates through web-based transactions. These are accessible through https://digitalid.verisign.com/services/ client/index.html. Verisign OCSP Responder Certificates are available through query of the Verisign LDAP directory server at directoy.verisign.com.

[39] [http://www.verisign.com/repository/CPS/]. Accessed Aug 2003.

[40] [http://www.verisign.com.au/whitepapers/enterprise/revocation/cert_revktoc.shtml]. Accessed Aug 2003.

- Managed PKI customers may contract for OCSP services; the URL information is then provided to the individual customer. Verisign also operates several Infrastructure CA's that issue Certificates to Verisign infrastructure components (e.g. OCSP Responders providing Certificate status information and Roaming Servers, which support the Verisign Roaming Service).

G. VALICERT CPS

Valicert merged with Tumbleweed Communications Corporation in February 2003. This merger led to a change in the direction of their business. Valicert is exiting the services end of PKI and has sold the CA and trusted roots to other companies. The new direction at Valicert is the development of product and software implementations that will enable PKI. To this end, Valicert has developed and fielded Enterprise Validation Authority (EVA) 4.6, which is an OCSP implementation. The Valicert CPS was still in existence at the time of this writing and was used for comparison with other industry CPSs[41].

41 [http://www.valicert.com/repository/]. Accessed Aug 2003.

THIS PAGE INTENTIONALLY LEFT BLANK

V. SECURITY-RELEVANT ATTRIBUTES OF CERTIFICATE REVOCATION MANAGEMENT SCHEMES

A. SECURITY RELEVANT METRICS

There are many environments in which the DoD PKI may be implemented; it might be a stateside shore based installation with connection to a dedicated revocation mechanism, a ship at sea with SIPRNET and NIPRNET limited by bandwidth or connectivity, an aircraft carrier with near dedicated SIPRNET and NIPRNET connectivity and dedicated revocation mechanism checking hardware, an advance team in a foreign country with unknown connectivity and no ability to maintain connectivity with revocation mechanisms, or any number of environments in which the DoD operates. It is clear that no sweeping statements or generalizations should be attempted regarding the environment in which a DoD PKI may be implemented. Instead of offering an ordering for each environment, this thesis generalizes across the band of environments by first developing the security relevant attributes that comprise all certificate revocation implementations, building a set of metrics based on this examination, offering a "best practices" security model for certificate revocation; and finally, reviewing individual implementations. While generalizations of environment are not offered here, it will be worthwhile to examine the requirements that guided the implementation of DoD PKI alongside the various revocation mechanisms. On August 2001, the National Security Agency released "Department of Defense Target Public Key Infrastructure Operational Requirements Document", this document established a minimum set of acceptable standards that a DoD PKI implementation should meet or exceed.

The primary research question of this thesis is to determine the security-relevant strengths and weaknesses of all currently operational, proposed, or theoretically possible implementations of certificate revocation management, and then provide a ranking of said implementations. This goal differentiates itself from ranking certificate revocation implementations by methods such as most economical, most feasible, or some other standard. As such, a set of metrics related to security need to be identified and then used to compare the relative merits of the differing revocation methods. An accepted method for examining an information security issue is to explore the attributes of confidentiality,

authenticity, integrity, and availability. It will be demonstrated that the attributes of confidentiality, authenticity, and integrity need not be integrated into the process of developing a set of security relevant metrics. It will be further demonstrated that the attribute of availability is the most critical security relevant attribute in a certificate revocation implementation. By examining availability in detail, a set of metrics will be developed. The end result of a secure revocation mechanism will be that the end user is able to quickly and reliably determine certificate status and have a greater confidence in the validity of the certificate in question.

In a Public Key Infrastructure there is no need to keep the certificate number of revoked certificates confidential; CRL's in all forms and current implementations are publicly available for viewing by any party. A primary goal of a certificate status mechanism is to provide the widest dissemination of the certificate status to interested parties. This dissemination strengthens the PKI by assuring the relying party that certificate validity may be ascertained and a determination made as to whether the certificate may be relied upon. This public nature of certificate status does pose a security risk; an actor that has stolen an individual's private key or an individual that has had his/her status as a trusted agent revoked may readily determine whether the certificate in their possession remains valid and might be used for unscrupulous purposes. To prevent this, it is possible that an organization might implement access controls or build mechanisms of confidentiality that limit information regarding certificate status; this would allow certificate status to be viewed only by authorized entities or individuals. This type of Public Key implementation is outside the bounds of the DoD PKI. For these reasons, all current, proposed and theoretical implementations of PKI certificate status mechanisms need not be ranked by confidentiality.

Integrity and Authenticity are each adequately provided in all current, proposed and theoretical implementations of certificate revocation mechanisms. Data integrity and authenticity are ensured in each method of certificate status checking by some form of digital signature across the data that is being relied upon. The specific methods in which these signatures are implemented differ for periodic publication mechanisms and online query mechanisms as discussed next.

Periodic publication mechanisms, be it a CRL, Delta-CRL, or other variant, are mechanisms which provide the known universe of revoked certificates under a particular issuing CA. This mechanism generally presents a digitally signed list of all previously valid certificates that have been revoked before their scheduled expiration date. The list contains both the time of its publication as well as the time of the next expected publication. The issuer's digital signature prevents the possibility of changing, deleting, or modifying the data without the relying party realizing the data is no longer valid. Inclusion of the publication times provides a mechanism through which the relying party may determine whether the list in their possession is current or has been superceded by a more current list.

An online query mechanism provides a certificate-specific response to a user's request regarding the status of a certificate. Currently OCSP is the only implementation--within the X.509 framework--that utilizes an online query mechanism. The attribute of authenticity could be attacked by a man-in-the-middle (MITM) replaying a previous OCSP response--which indicated that the certificate in question is valid--to the relying party. This threat is removed by the use of nonces. In OCSP requests, the nonce is identified in the *requestExtension*, while in OCSP responses the nonce is identified in the *responseExenstions*. The nonces implement a response validity interval. The use of nonces "cryptographically binds a request and a response to prevent replay attacks".[42] The entire response is signed by "either the CA that issued the certificate whose status is being checked, by a responder that has been authorized by the CA, or by a responder that is trusted by the requester."[43] Because the attributes of Integrity and Authenticity are not factors which affect any of the current, proposed, or theoretical certificate revocation mechanisms, there is no need to derive a set of metrics for security order ranking based upon these attributes.

It has been demonstrated that three of the attributes of information security--confidentiality, authenticity, and integrity--are not factors that affect the security ordering of certificate revocation mechanisms. Availability is the principle attribute against which

[42] RFC 2560, PKIX OCSP, para 4.4.1.

[43] Concept of Operations for the Department of Defense Online Certificate Status Protocol Service, Electrosoft Services Inc., November 12, 2002.

differing methods of certificate status checking need to be measured. Evaluation of each of the differing methods of certificate status checking against metrics derived from the attribute of availability will facilitate a security ordering of the implementations. When a user is presented with a certificate, he or she must have a means for determining the status regarding revocation. Without this mechanism, the authenticity of the certificate must be called into question. Differing certificate status checking mechanisms have differing ways of ensuring that a relying party may trust the status of the certificate in question. Dr. Andrew Nash of RSA security offered this opinion: "Availability is clearly the most significant aspect of dealing with any public key system where the identity is important…In the case of an online validation scheme, access to the validation service or responder is key and is one of the more important design considerations." Dr. John Hines of Valicert Technology echoed this sentiment: "we put a great emphasis on being highly available and work with load balancers and hardware signing modules."

The attribute of availability needs to be distilled into elemental parts in order to define a set of metrics. The Committee on National Security Systems defines availability as "timely, reliable access to data and information services for authorized users."[44] It is tempting to state that this definition equates to "availability of revocation information upon demand by the relying party." However, this definition is excessively broad for the purpose of extrapolating metrics for the purpose of assessing a revocation status mechanism's security strength. A more precise definition of availability with regard to certificate status would include the following elemental parts:

- Notification Latency: Time between which a certificate is deemed to require revocation and notification of such is delivered to the CA for inclusion into the next publication of the CRL.
- Publication Latency: Time between notification of a revocation and publishing of said information to a publicly accessible location.
- Query-Response Latency: Time between which a user issues a query and subsequently receives the corresponding response.
- Sending party vs. Relying party revocation responsibility: Whether the onus to perform the processing associated with the revocation check is on the sender/owner, receiver/user, or both.

44 [http://www.nstissc.gov/html/library.html]. Accessed Sep 2003.

- Update Granularity: Whether the entirety of the information is updated or whether it is updated through an incremental or differential scheme.
- Push or Pull: The method through which the revocation information is disseminated; i.e., is the information pushed to the relying party or does the relying party actively pull the information from a repository.

The primary parts of the definition that will be considered for use as metrics are publication latency and query-response latency. The remaining parts, notification latency, sending party vs. relying party revocation responsibility, update granularity and push/pull methods, are useful for precisely defining availability, yet they do not allow for any greater specificity of security in terms of certificate revocation mechanisms. The rationale for discounting these parts of the definition is as follows:

Though Notification Latency contributes to the total amount of time until the relying party may discover that a certificate has been revoked, the notification phase is ultimately reliant on user discovery of revocation requirement and the ensuing activation of the revocation process. This time factor of user intervention may be affected by policy but it is not readily affected by changing the implementation of the PKI model. Therefore, it is not a metric that applies to this discussion of security of certificate revocation implementations.

When considering sending party vs. relying party revocation responsibility, it is assumed that whether the onus to perform the computations required to check the status of a certificate is on the owner of the certificate or the relying party, that any device implementing X.509 revocation checking will have sufficient processing power in order to complete the required computations for this process.

Update Granularity – The operational environment the PKI is implemented in will mandate what minimum and maximum levels of granularity are acceptable. This is a function of policy and is mandated by the CP and CPS.

Push/Pull methods – The method in which revocation information is received, whether the transmission of data is initiated by the CA or by a relying party, is of no consequence. The availability of data may be made the same through specifications in the CP and CPS.

Each of the preceding four parts of the availability definition applies equally across all PKI environments and may thus be considered negligible for use as a metric. While these parts assist in defining the broader definition of availability, they do not add clarity in terms of classifying certificate revocation methods in a security ordering.

The dissection and classification of the final two parts of the definition of availability facilitates building security-relevant metrics. It is reasonable to assert that "currency of revocation information" is a measure of the degree to which data presented to a relying party accurately reflects the true status of the data. The two factors whose sum equals the whole of "currency of revocation information" are notification latency and publication latency. Notification is a function of organizational policy and operational implementation; it is not required for evaluating validation mechanisms based upon currency and timeliness for the reasons discussed above. Publication latency may be affected by both the mechanisms and policies implemented by the CA. Publication latency is relevant to the discussion of certificate status validation and is directly linked to the "currency of revocation information". From this point on publication latency, which implies "currency of revocation information" for our use in this thesis, will be referred to simply as currency.

In order to use currency as a metric, a group of time periods must be identified. The generally accepted update frequency for certificate revocation information is once every twenty-four hours, this is also the requirement as listed in the DoD Target Public Key Infrastructure Operational Requirements Document. The update of once every twenty-four hours is assigned as the middle of three currency groups: High, Medium, and Low. From this start point, the following currency values are developed: High= Eight hours or less, Medium= 24 hours, Low= greater than 24 hours. A higher currency should equate to higher degree of trust in the validity of all certificates. When currency is high, the amount of time between when the CA administratively verifies a certificate is in need of revocation and the time this information is available to the relying party is very small. This provides a small window of opportunity for misuse by a party that had obtained a certificate for which they did not have authority. Low currency would indicate a greater amount of time available for misuse before the information was available to the relying

party; this would engender lower confidence in the PKI in general. For the purpose of evaluating certificate status checking, it can be stated that currency is a function of publication latency, or, Currency = f{Publication Latency}.

It can also be stated that, for the purposes of certificate status checking evaluation, availability is a function of "query-response latency", or, Availability = f{Query-Response Latency}. It may be argued that currency of information is implied by availability; however, because currency is not sufficiently explicit in the definition of availability it should be considered separately and used as a metric for determining the security strength of a given certificate validation implementation. The term query-response latency is not precise enough to use as a metric and is more fully defined by the term Revocation Information Availability (RIA). Revocation Information Availability (RIA) is comprised of four constituent factors:

- The ***size of the response data structure*** used during the certificate status checking transaction. The size of the response is dependent on the mechanism employed; an online mechanism would include the transaction size of both the request and response along with any incidental communications while a periodic publication mechanism such as a CRL, Delta-CRL or Indirect CRL would include the download of any portion of the list. Both methods employ some amount of nominal overhead for initiating the validation response; the distinction is in the size of the response transmitted to the relying party. An OCSP type method will transmit a small signed response while a CRL method will transmit either an entire CRL or some portion thereof.
- The ***bandwidth profile*** that is a measure of the network connectivity between the relying party and the certificate status checking transaction. The general definition of bandwidth is the amount of data that is carried over a medium, while throughput is the amount of data carried over the same medium over some period of time. Revocation mechanisms may be characterized as being tolerant of high or low bandwidth environments; another way of stating this is that they may operate in a "Thin Pipe" or "Fat Pipe" environment. This definition is not complete, while each implementation may have a preferred environment; the function of bandwidth consistency has not been addressed. The mechanisms must be measured against network consistency, designated as steady state or intermittent, to determine their resiliency in each state. It should be apparent that an ideal environment would have a bandwidth profile characterized by high bandwidth and steady state connectivity; however, it is a possibility that some of the environments in which a DoD PKI is implemented may have a varying degree of each of these factors.

- ***Response Generation Latency*** is a measure of the time from the request for validation to the time that the response has been computed and is ready for transmission to the relying party. Revocation mechanisms that entail resource intensive processing should rely upon the validation server to perform the processing. This strategy is more conducive to the typical (or worst-case) scenario of a "thin" client. Depending on the type of implementation, the data may be sent from a central point to the relying party or there may be no transmission if the data resides locally on the computer that generated the request. These two implementations are differentiated by the location of where the computations are completed; this is the classic "Thin Client" vs. "Fat Client" model.
- ***Proximity*** of revocation information in relation to the relying party may be further divided into three components. The first, network distance, is somewhat indeterminate. It may be the physical distance from the relying party to the mechanism providing a response, the number of hops required for the transaction to be completed, or some combination thereof. The second, degree of query forwarding or handling, incorporates the concept of online forwarding of requests or pointers to periodic publication mechanisms. The final component, repository redundancy, incorporates the concept of storing the information in multiple locations. A method that has the same information stored in a distributed format may be able to generate faster responses based upon the previously mentioned metric of network distance, and may therefore be able to guarantee a more timely response to the relying party

For the purposes of evaluating the security of certificate validation mechanisms, Currency and Availability have been identified as primary metrics. These metrics have been distilled to their component parts where the validation mechanism availability is a function of currency and availability [Validation Mechanism Availability = f(Currency + Avail)]. For the purposes of evaluating certificate validation mechanisms, currency is defined as a function of publication latency [Currency = f(pub-latency)] and availability is a function of response size, bandwidth profile, response generation latency, and proximity [Avail = f(response size + bandwidth profile + response generation latency + proximity)]. The availability metrics are graphically depicted in Figure 12. The possible environments which may arise will necessarily encompass different values for both currency and availability; this can be visualized as a continuum as represented in Figure 13; the differing values for currency and availability will define an end state in a particular quadrant.

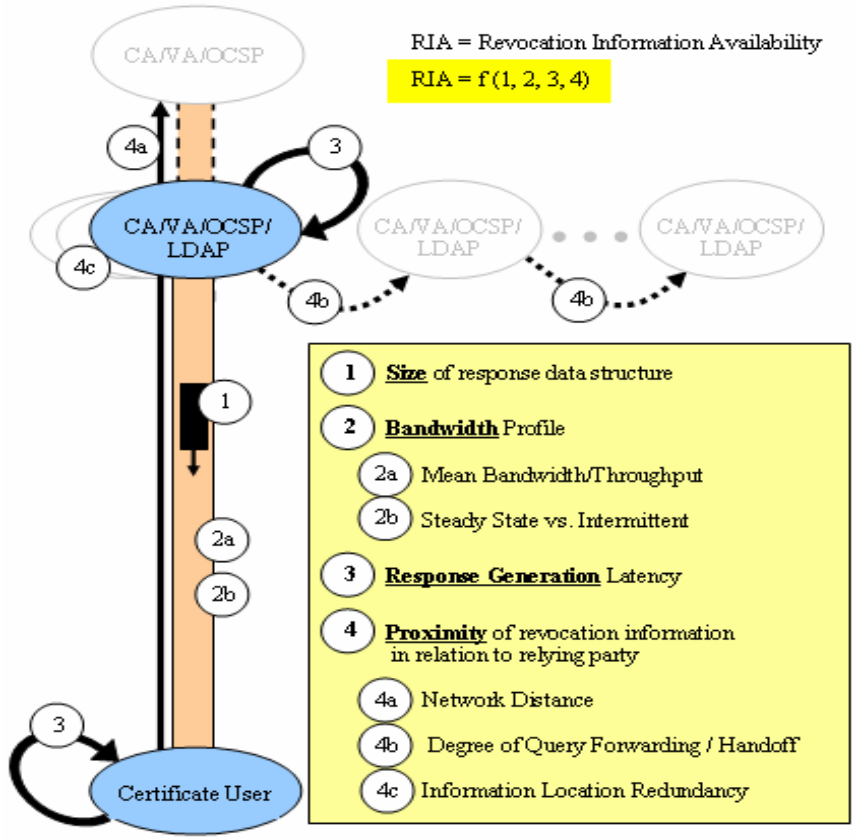

Figure 12. Model for Revocation Information Availability (Query-Response) Metrics.

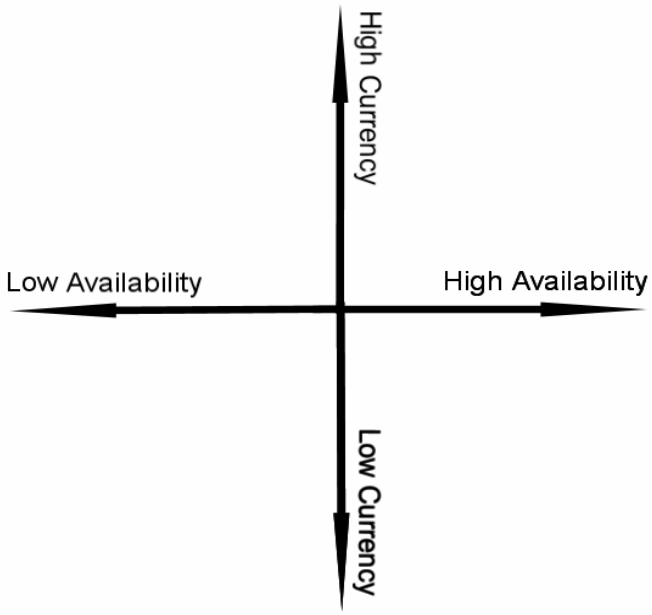

Figure 13. Implementation Continuum.

Before building a "best practices" model, it is worthwhile to consider the set of acceptable standards regarding currency and availability as set forth in the "Department of Defense Target Public Key Infrastructure Operational Requirements Document." Several pertinent issues are noted:

- …system performance parameters are different depending on the environment in which the PKI is being used. Most notably, compromise notification response time is much more critical in a tactical environment. Although it is recognized that these differences exist, functionally the PKI being procured does not differ based on the environment to which it will be deployed. Other than revocation performance parameters a PKI deployed in a tactical environment has no special functional requirements. The effect the environment has on the PKI will be on how it is implemented.
- PKI subscribers need to use their certificates with any DoD computer/platform, regardless of the operating system of the computer. PKI must supply…certificates that are, to the maximum extent possible, computer platform and operating system independent.
- PKI must allow subscribers to invoke certificate revocation…without requiring direct, local contact between the subscriber and the CA or RA who performed the initial registration of that subscriber.
- PKI must make certificate revocation information widely available so that supported applications can determine the validity of individual certificates before relying on them.
- PKI must comply with the requirements of the DoD CP regarding CRL's.
- PKI must be able to publish CRL's in forms that take into account the constrained communications and directory access capabilities of some supported applications.
- For some supported applications, CRL's will be an inappropriate or ineffective means of promulgating revocation information. PKI must support alternative means for supported applications to check certificate validity via an on-line, interactive capability. This capability must be based on commercial standards.

In addition to the preceding statements, the publication latency requirements are defined as both a threshold and an objective level for either Tactical or Non-tactical environments. These requirements are summarized in Figure 14.

Level	Non-tactical	Tactical
Threshold	Once a day and within 6 hours of comprimise notice being recieved by CA	Once a day and wihtin 1 hour of comprimise being received by CA, for comprimises withing the tactical area of operations
Objective	Once a day and within 15 minutes of comprimise notice being recieved by CA	Once a day and wihtin 10 minutes of comprimise being received by CA, for comprimises withing the tactical area of operations

Figure 14. Publication Latency Requirements[45].

Requirements for tactical networks further states that "PKI for tactical networks must be able to provide continuous support despite interruptions in communications to fixed networks. PKI support to connected tactical networks must provide a full range of capabilities to….revoke certificates….and disseminate PKI –generated objects…. PKI must be able to provide revocation information in a condensed form to minimize the communications load on a limited tactical communications circuit. PKI supporting tactical networks must be able to locally (i.e., in theatre) generate and disseminate revocation information of interest to the tactical area of operations." The maximum non-availability period of an individual CA within a seven-day period must be no greater than four hours; requirements for CA availability relevant to revocation functions are summarized in Figure 15.

CA Availability Requirements	
Threshold	Objective
Single CA 168 hours per 164 hour period of availability Operationally available rate of 99.9% No single points of failure that would make CA services unavailable	Single CA 168 hours per 168 hour period of availability Operationally available rate of 99.99% PKI msut be able to restore any non-available services within 4 hours of a non-catostrophicfailure

Figure 15. CA Availability Requirements[46].

[45] Department of Defense Target Public Key Infrastructure Operational Requirements Document, p. 28.

[46] Ibid.

The "Department of Defense Target Public Key Infrastructure Operational Requirements Document" makes clear that a high premium is placed on the attributes of availability.

From the previously developed metrics, the implementation continuum and the goals set for DoD PKI, a "best case" PKI implementation model may be proposed. This model will be used in conjunction with the associated metrics of availability and currency to rank all currently operational, proposed, or theoretically possible implementations of certificate revocation management. The proposed "best case" PKI implementation model implements mechanisms which provide high currency of data coupled with high availability while remaining tolerant of the environment. The proposed "best case" PKI implementation model is described thusly.

Using an over-issued CRL scheme, the CA publishes complete lists of revoked certificates twice daily to multiple external LDAP directories available throughout the PKI environment; these updates are available for immediate download (pull) by online validation response implementations. There are multiple authoritative validation response servers to which the relying party may connect; at least one of these servers is geographically or logically collocated within the end users organization while other authoritative validation response servers are located outside of the organization. Each time the validity of a certificate is called into question, a real time query regarding the status of the certificate is made. Certificates published by the CA point to the primary and alternate responders; the relying party uses an online revocation checking mechanism and is able to connect to any of the validation responders specified in the certificate. The responsible CA publishes a daily full CRL as well as hourly Delta-CRL updates to each previously issued CRL. A mirroring LDAP resides within the physical or logical boundary of the organization to which the relying party belongs; the mirrored CRL information on the local LDAP is the primary repository for information for the local validation responder. In the event, that the bandwidth profile does not sustain downloads of full or Delta-CRLs to the local LDAP, the primary validation responder seeks to work up a hierarchical chain seeking current information. In the event that communication cannot be established with the local validation responder, the PKI mechanism seeks to establish communications with a validation responder higher in the hierarchical chain.

The LDAP located within the physical or logical boundary of the relying party maintains a local cache of the information up to the time at which connectivity to the primary LDAP servers was lost, or the bandwidth profile was reduced to the point of being unable to download the CRL or Delta-CRL. Because the organizational LDAP is within the physical or logical boundary of the relying party, it is assumed that there are no bandwidth constraints between the relying party and this server. It should be noted that when communication can neither be established nor maintained with either the local or remote validation responders, no new certificates to be verified will be received either. A further assumption is that notification regarding the status of the certificate will always be made to the relying party, if certificate status cannot be obtained, specific information regarding why this information was not obtained will be displayed to the relying party. Additionally, a "CRL grace period"[47] is implemented. During this period, a mechanism will report to the user what the time of the last connection to the primary servers was and what the status of the certificate was at that time. Based upon the provided information, the relying party will be able to make an informed decision regarding the validity of the certificate based upon the specific information provided. This decision will be one of three options:

- Reject the certificate due to insufficiently low currency.
- Accept the certificate during the grace period. The user relies upon the last known status of the certificate at a known time.
- Accept the certificate based upon operational necessity as dictated by higher guidance/authority.

The proposed "best case" implementation allows the relying party to use centralized processing in low bandwidth environments. Although developed independently of the model proposed in the OCSP Concept of Operations document;[48] the two models share a great deal of similarity. The OCSP Concept of Operations document sought to find a workable implementation of OCSP, while this thesis seeks to

[47] Concept of Operations for the Department of Defense Online Certificate Status Protocol Service, Electrosoft Services Inc. November 12, 2002, p. 7.

[48] Concept of Operations for the Department of Defense Online Certificate Status Protocol Service, Electrosoft Services Inc. November 12, 2002.

propose a "best case" security implementation. Both efforts arrived at the same conclusion for a proposed model. The OCSP Concept of Operations document outlines the following benefits of this type of implementation:

- Reduction of traffic related to certificate validation. The relying party does not have to download an entire CRL, therefore the amount of bandwidth required to process both each individual as well as the total number of transactions is reduced.
- A higher level of reliance may be established; the relying party can be ensured that the most recent data has been published and is available for revocation purposes.
- Bandwidth constrained environments are able to perform revocation checking.
- An architecture which is both highly redundant as well as scalable across the DoD.

The following sections review the specific revocation mechanisms, enumerated in Chapter III, in terms of strengths and weaknesses as per the continuum of availability and currency and evaluated through the previously described metrics that are listed in Figure 16.

			Bandwidth Profile			Proximity		
	Currency	Data Structure Response Size	Mean Bandwidth / Throughput	Steady State vs Intermittent	Response Generation Latency	Network Distance	Degree of Forwarding / Handoff	Information Location Redundancy
Implementation								

Figure 16. Evaluation Matrix

B. REVOCATION MECHANISM RANKED BY SECURITY RELEVANT METRICS

1. Ranking Methodology

Each mechanism will be evaluated using the previously developed security metrics. In many instances, there are several methods in which the mechanism may be implemented. Specific mention is made of the differing types of possible implementations; however, only the most clearly robust method of implementation is ranked. In cases where it is not clear which method is the most robust, all implementations are ranked. Each mechanism is reviewed individually using the pertinent security relevant attributes and given one of the following rankings: Above average, Average, Below average, or Does not apply. The methodology for assigning a ranking is as follows:

Above average – a mechanism in the implementation directly addresses the security ordering metric in order to maximize performance in relation to the metric; the performance is significantly superior to the other implementations reviewed.

Average – no mechanisms in the implementation address the security ordering metric in order to maximize performance in relation to the metric; the performance is on par with the other implementations reviewed.

Below average - no mechanisms in the implementation address the security ordering metric in order to maximize performance in relation to the metric; the performance is significantly inferior to the other implementations reviewed.

2. Proposed "Best Case" Mechanism

It is the contention of this thesis that the most secure method for providing certificate validity information is through the proposed "best case" certificate validation mechanism.

a. Currency

Currency for this model is evaluated as above average. The best case model has two methods that give it above average currency, both of which would be mandated by the CP and CPS. The first is the requirement for publication of a complete list of certificates on a twice-daily basis. The second is the mandate for hourly Delta-CRL updates. In effect, the oldest information that a relying user would have when the

system is functioning correctly is 59 minutes old. Most PKI's, including the DoD PKI, require updates of only every 24 hours. This indicates that the best case model will be one of the few to have an above average currency.

b. Data Structure Response Size

The data structure response size is inversely proportional to its rating, a system with a smaller response size is evaluated as above average while the larger the size grows the more negative the rating becomes. In this implementation, the data structure response size is rated as above average. The response to a request for certificate validation is similar in structure to a signed OCSP response and would comprise roughly 2000-4000 bytes, depending upon key size.[49]

c. Bandwidth Profile

The proposed best case model is tolerant of both low bandwidth and intermittent connection status; it therefore is rated as above average in both of these categories. This tolerance is achieved by several mechanisms. Use of an online validation response scheme allows tolerance for low bandwidth environments. When bandwidth is available, the complete CRL as well as any delta CRL updates are downloaded to the LDAP server located within the physical or logical limits of the organization. This mirror provides the ability for the end user to receive an online OCSP response in the event that a connection cannot be established with the validator in an upper level hierarchy. In the event that bandwidth does not allow a real time validation response from the upper level validator, and no CRL has been downloaded within 24 hours, the CRL grace period information is displayed and allows for an informed decision regarding use of the certificate.

d. Response Generation Latency

The measure of the time from the request for validation to the time that the response has been computed and is ready for transmission to the relying party is inversely proportional to its rating; the more quickly a response can be generated the higher the rating. In the proposed best case model, the computation of a response may be performed at any number of validation authority locations and is not based upon the processing power of the relying party. Since it is generally accepted that a validation authority

[49] [http://www.corestreet.com/whitepapers/CertificateValidationChoices]. Accessed Nov 2003.

server will be specialized, and therefore optimized for such purposes; whereas a relying party is more likely to be a general purpose computer, it is logical to rate the response generation latency as above average.

e. Proximity

This model has the ability to adapt to changing network conditions. The distance from the relying party to the validity responder is dependent on the bandwidth profile at the time of validation request. When the bandwidth profile provides a steady state medium, the relying party will process validation requests through the local responder. When the bandwidth profile is intermittent the relying party will continue to process validation requests through a local responder as long as the local responder has the most recent copy of the full and Delta CRL's. When bandwidth profile is such that the most recent copy of the full and Delta CRL's have not been obtained by the local responder, the relying party is redirected to an authoritative responder higher in the hierarchical chain. This transaction between the relying party and the remote responder is tolerant of low bandwidth environments; only a request and response are exchanged vice a download of information. The ability to adapt to changing network conditions ensures that the physical or logical distance between the responder and relying party is less of a factor. There is no query forwarding or handoff up the validation chain that would add size to the request or increase the time before a request could be received; the "down-chain" fall back is implemented by design. In the event that the authoritative validation responder cannot be reached, the secondary or local validation responder contains mirrored information for complete repository redundancy. For these reasons, all factors associated with proximity are deemed above average.

3. Certificate Revocation Lists (CRL)

The primary disadvantage of the CRL mechanism is that a complete CRL must be obtained and processed in order to verify the status of a single certificate. The CRL grows in direct proportion to the number of revoked certificates; revoked certificates must be kept on the list as long as the certificates have not expired. As stated earlier, each revoked certificate on the list occupies approximately 9 bytes of data. A CA with

thousands of end users would be required to provide a mechanism to push or pull several megabytes of data to provide for revocation checking needs. The Complete CRL has several undesirable security implications.

a. Currency

The data within the CRL is only valid for the period of time as defined in the CP, which in practice is generally twenty-four hours. The data remains valid until the *NextUpdate* time; at which time an entire new CRL must be downloaded. This meets the criteria for medium currency and a rating of average.

b. Data Structure Response Size

The data structure response size is equal to the size of the entire CRL. This response may be megabytes in size. The complete CRL may be distributed to several sites, cached on the organizational network of the relying party, or pushed to the relying party on a routine basis. Regardless of location, the end user will be required to download the CRL when the *NextUpdate* time is reached. This is one of the primary disadvantages of the standard CRL and is evaluated as below average.

c. Bandwidth Profile

In order for a large response to reach the relying party, the network must possess a large bandwidth to facilitate retrieval or pushing of the CRL from the CA and authoritative repository and must necessarily be steady state to ensure the entirety of the download is received without corruption. In large networks where users frequently check the status of certificates, each workstation or user may be required to download a new CRL often; the bandwidth required to support these downloads--even within the limited scope of the relying parties intranet--may be prohibitively large. This CRL is not tolerant of intermittent or low bandwidth and is evaluated as below average.

d. Response Generation Latency

The generation of responses in reply to a certificate validation request is completed by the relying party using the computational power of the machine that hosts the full CRL; in the case of a full CRL the relying machine need only transmit the data (CRL) to the relying party or a responder acting on its behalf. The computational power required to verify a signature attached to a CRL, parse the data, and make a judgment regarding the validity of the certificate based on this information is relatively trivial.

Most current desktop or laptop computers and handhelds such as Blackberry's, PDA's, or cell phones possess the requisite processing power required to complete this task. This is evaluated as above average because the amount of work to be completed and the time involved is relatively insignificant.

e. Proximity

The distance in which a full CRL must traverse from a directory to the relying party computer hard drive depends on the specific implementation. If only available for download from the authoritative repository (ie the CA repository) then the distance, both physically and logically, may be considerable. In the full CRL model, both network distance and repository redundancy are closely coupled; implementation specific designs may allow for downloading from the authoritative location, followed by storage of the CRL in several locations closer to the relying users' organizations. Implementations which store revocation information in a distributed format significantly increase information redundancy; the attribute of network distance may also be optimized by careful selection of where the information repositories reside. With this in mind, it is important to note that verifiers must be able to obtain up-to-date CRL's from every supported CA. Within the DoD framework this is relatively trivial, however in other PKI systems certificate chaining may make this difficult as there may be a large number of CA's and multiple CA's per certificate chain.[50] These factors earn both metrics a rating of average. There is no query forwarding or handoff in the full CRL implementation.

4. CRL Distribution Points or Partitioned CRL's

CRL Distribution Points are functionally similar to the full CRL. The metrics which require reexamination are query forwarding, size, and degree of handoff. Because the static CRL distribution points may be defined or partitioned, there is a greater amount of forwarding in the distribution point and partitioned CRL revocation methods. This scheme was designed to reduce the storage problem for the CA associated with a growing CRL; it also may be implemented by logical groups so that the relying party need download only a smaller subset of the entire CRL. While the overall size may be smaller, the size is still significantly larger than online implementations. This scheme has a

[50] [http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=/netahtml/search-adv.htm&r=2&f=G&l=50&d=ptxt&S1=valicert&OS=valicert&RS=valicert]. Accessed Nov 2003.

higher level of query forwarding and handoff which may also equate to a longer response time between the relying party and the validation authority. The degree of query forwarding/handoff is evaluated as average while the size is evaluated as average.

5. Redirect CRL's and Enhanced CRL Distribution Points

The Redirect CRL is designed to reduce the storage problem within the CA LDAP created by static partition points in standard CRL Distribution Points and Partitioned CRL's. By using dynamic partition points, the mechanism acting on behalf of the relying party may download only that portion of the CRL which pertains to the user. These partitions may be allocated by specific user groups, functional groups, revocation reasons, or numerical segmentation of the CRL. These different partitions may also be updated at different rates depending on the needs of the relying organizations. The segmentation may reduce unneeded downloading of entire sections of the CRL if the end user application is aware of the segmentation scheme and is able to both selectively download the needed CRL partition as well as verifying the validity period of each CRL partition before downloading. The evaluation of this scheme mirrors that of the CRL Distribution Points or Partitioned CRL's with one exception, data structure response size. When used with an end user application aware of the partitioning scheme, the Redirect CRL's and Enhanced CRL Distribution Points may significantly reduce the size of the response required for download by the end user. While the size is reduced, it still remains significantly larger than many of the online revocation checking mechanisms and is thus evaluated as average.

6. Delta Certificate Revocation List

The Delta CRL, issued between two CRL updates, is similar to the CRL; however, the implementation of updates changes several key metric rankings.

a. Currency

Governed by the CP and CPS, the currency of the information available to the relying party may be increased by more frequent Delta CRL updates relative to the full update of the CRL. This type of implementation is evaluated as above average.

b. Data Structure Response Size and Bandwidth Profile

There are two components to the Delta CRL implementation, the full CRL and the Delta CRL which includes the updates which occurred since the last full CRL

was issued. Users need only download a new full CRL when it is issued and are able to download the smaller incremental updates when necessary in the interim. Depending on how this is implemented (e.g., an organization which has a weekly full CRL coupled with daily Delta CRL's) this may significantly reduce the size of the data structure that is transmitted. An implementation using this type of update would also have a significant impact on the bandwidth profile. By using Delta CRL's the requirement for multiple users in an organization to download multiple instances of the same complete CRL daily is negated; the users need only retrieve updates (deltas) to the base (full CRL). This has a significant impact on the bandwidth required to make validity information available to the end user. If configured to download full CRL's during times of historically low traffic volume and then rely on the smaller delta updates during intervals of higher volume this implementation would be more tolerant of a low bandwidth environments. This type of implementation is evaluated as above average for data structure response size and above average for bandwidth profile/throughput.

7. Indirect Delta CRL's

The Indirect Delta allows for simplified retrieval of certificates from several different CA's by maintaining lists signed by several entities in one partition. This model is similar to the full CRL except that it reduces the degree of forwarding or handoff required for certificates to be used in the cross trust environment; it is evaluated as average in this area. By using an Indirect Delta CRL's the likelihood that the relying party will be able to access certificate information germane to his certificate is increased. An ancillary benefit of this implementation occurs when many PKI's implement Indirect Delta CRL's; the likelihood is that the network distance will be decreased. However, as this is not necessarily controlled by the end user organization, it does not change the rating for network distance.

8. Sliding Window Delta Certificate Revocation Lists

The sliding window Delta CRL's, as with all Delta CRL implementations, is designed to provide efficiency gains in administering the CRL. Specifically, the sliding window Delta CRL and over-issued CRL's seeks to reduce the peak rate of requests for new CRL's and force a more even distribution instead of having periods of peak intensity following the *NextUpdate* time. This performance can be directly tied to the bandwidth

profile requirement. This allows for the use of a lower overall average bandwidth/throughput because the peak request rate is decreased. This mechanism is evaluated as above average in mean bandwidth/throughput while retaining the same remaining metric ratings as the Delta CRL method.

9. Certificate Revocation Trees

The goal of the CRT to "greatly reduce the processing effort, network bandwidth, network latency, data storage, and data replication requirements needed to determine whether a particular certificate has been revoked….allow(ing) certificate status to be determined without knowledge of the entire list of revoked certificates and without having to search the entire list of revoked certificates..."[51] directly impacts several security metrics. An attribute not directly measured by the response generation metric relates to the amount of work required to compute the CRT. The trees are regenerated every time a new revoked certificate is added to the leaf nodes of the tree. This new structure only requires re-computation of the sub-tree. Processing (work) done by the CA server computing the tree may be done asynchronously, that is, it is computed without regard to queries. The CRT may be implemented in an online responder fashion or as a periodic publication mechanism. The CA need only sign the CRT one time and no public key must be online, unlike the OCSP model where a key must be online and each response must be signed.

a. Currency

The currency the CRT method is defined by the CP and CPS. Each time the raw (i.e., un-hashed plaintext) CRL is updated as newly revoked certificates are added to the list, the branches affected by the new leafs must be recomputed; this does not necessitate re-computation of the entire tree. While the amount of work required to respond to certificate requests is small and may be handled by a XML enabled responder, the amount of work to be done in re-computing the hash tree is significant. While no data on this work appears to be in the public domain, it is inferred that this work would be the limiting factor in publication latency. Naor and Nissm built on the original Merkle tree and suggested using a 2-3 tree in which every interior node has 2 or 3 children and the

[51] Patent 6442689; US Patent Office.

paths from the root to the leaves always have the same length.[52] This allows for changes to the tree where the insertion and deletion of an element affect only the nodes on the insertion/deletion path. The time required to complete this operation is proportional to the height of the tree, which is proportional to the base-2 log of the number of leaves (revoked certificates) on the tree. Naor and Nissim state that the "CA-to-directory communication costs…are optimal (proportional to the number of changes in the revocation list), enabling high update rates". This is evaluated as above average.

b. Data Structure Response Size

The data response size of the CRT is much smaller than that of the CRL but larger than that of an OCSP response. As previously stated, the CRL growth is directly proportional to the size of the PKI and the number of revoked certificates. The CRT grows logarithmically in response to the number of revoked certificates.[53] A doubling of the number of certificates revoked doubles the size of a CRL; the CRT meanwhile only grows as a base-2 log of that number. The response size depends on the location of the requested certificate within the tree and is based upon the number of leaves and branches that need to be transmitted to provide the relying party with a proof. This is evaluated as above average.

c. Bandwidth Profile

The small size of the CRT response makes it tolerant of low bandwidth environments. Although not explicitly stated in any documents describing the implementation of this mechanism, there is no reason that this method could not be used in a distributed method to allow for tolerance of intermittent bandwidth environments. Implementation of responders is similar to that of distributing a CRL; there is no need to keep the information secret as all the data has been hashed and signed. Both the metrics are evaluated as above average.

d. Response Generation Latency

The response generation latency is two fold. The responsibility for replying to the request with the appropriate leaf, branch and root node information is

---

[52] Moni Naor and Kobbi Nissim. *Certificate Revocation and Certificate Update*. In Proceedings of the 7th USENIX Security Symposium, 1998.

[53] Shimshon Berkovits, Herzog, Jonathan. A Comparison of Certificate Validation Methods for use in a We b Environment. Mass: Mitre Technical Report, November 1988. Par 2.4.

relatively trivial. The relying party must take the received information and compute hashes to verify the proof, only then is the response generation complete. Similar to the full CRL, most current desktop or laptop computers have adequate processing power to complete this task. It is possible that smaller handheld devices such as blackberry's, PDA's, or cell phones may not be computationally robust enough to complete the proof required to check the status of the certificate. This is evaluated as below average.

e. Proximity

As previously noted, there is no reason that network distance and repository redundancy cannot be enhanced; the CRT method does not require a secure directory nor does it require a private key on the responding machine. This is evaluated as average for network distance and repository redundancy; there no query forwarding in the CRT implementation.

10. Trusted Directories

An evaluation of trusted directories is not included in this document because of their limited utility in a larger organization. The trusted directory, and therefore the entire PKI, is only as strong as the intranet on which it resides. To extend the PKI beyond the organizational level, the trusted directory must be accessible to parties outside of the organizational boundaries. Opening the trusted directories further weakens the security of this implementation, and keeping the directory from external relying parties limits the scalability of the implementation.

11. OCSP

OCSP responders may be implemented in several ways, through a Direct Trust, Delegated Trust, or Self-Signed Trust implementation. Each requires that some private key be located on the server to sign responses to validation requests. While not a metric, it is worth noting that physical security must be maintained on each server. Another vulnerability not directly addressed by the metrics is that revocation status information is signed; however, error messages sent by the OCSP responder are not.[54] A MITM attacker could attempt to confuse the relying party by sending false error messages.

[54] Baltimore Technology. The Online Certificate Status Protocol. [http://www.baltimore.com/devzone/pki/ocsp.asp]. Accessed Jul 2003.

a. Currency

Most OCSP responders use the full CRL or Delta CRL to obtain certificate revocation information.[55] Some OCSP implementations, such as Valicert's validation authority, pull the CRL information from several repositories or vendors allowing for cross certification. This information therefore is no more current than the standard CRL information located within the repository. Changes to this currency may be mandated through changes to the CP and CPS. The current implementations are rated as above average.

b. Data Structure Response Size

Compared to the CRL methods, the OCSP response size is small. The response to a validation request is either "good", "revoked" or "unknown" which is then signed by the CA[56]. Depending on the type of signature based scheme used, the OCSP CA signature is a minimum of 2,048 bits added to several bits for the actual response. Comprising several thousand bytes, the OCSP data responses are the smallest of any examined by this paper. This is rated as above average

c. Bandwidth Profile

OCSP has two distinct bandwidth environments. The first is from the OCSP responder to the CA or LDAP directory to retrieve the CRL information that is used by the responder. The second is the environment between the responder and the relying party. The environment between the OCSP responder and the entity with the CRL information to be downloaded is virtually indistinct from that of a relying party downloading a full CRL. In low bandwidth environments, a responder will be unable to access the CRL and will be unable to provide information to the relying party. The transaction between the responder and the relying party is lightweight in nature and is inherently tolerant of low bandwidth environments. The obvious drawbacks are that the responder must have sufficient bandwidth to both answer queries and download the full CRL's. This implies that the OCSP responder is not tolerant of low bandwidth environments when the *NextUpdate* time occurs. The OCSP implementation is

---

[55] Baltimore Technology. The Online Certificate Status Protocol. [http://www.baltimore.com/devzone/pki/ocsp.asp]. Accessed Jul 2003.

[56] Micali, Silvio. NOVOMODO, Scalable Certificate Validation and Simplified PKI Management. Massachusetts, MIT 199, p. 16.

vulnerable in the event that a denial of service is launched against the responder. In fact, an OCSP implementation which has too few centralized responders may be subject to an inadvertent denial of service when a large number of honest users attempt to access the responder and overload its capabilities. While there are many possible ways to implement OCSP; the implementations must be online to provide responses on demand to requests for validation. This mechanism is resilient to low bandwidth environments and is evaluated above average; its requirement that systems must always be online to provide responses, along with the inherent vulnerabilities to denial of service attacks rate a ranking of average for its low tolerance for intermittent bandwidth.

d. Response Generation Latency

The responsibility of response generation is the onus of the online responders. The work to build a response includes checking the status of the certificate, formulating a response, and signing the response before transmitting it to the relying party. While the time to compute the responses is relatively small, it is a CPU intensive operation, similar to that of building a CRT. Response generation is generally short but may become an issue when a large number of requests are received, whether they are comprised of valid user requests or denial of service attempts. Response generation latency is rated as average.

e. Proximity

The metrics of network distance and repository redundancy are directly affected by which type of OCSP responder is used. Because each server must have a copy of a private key to sign responses, the physical/logical distance from the relying party and amount of redundancy are directly correlated to the amount of resources the CA and relying organizations are willing to expend to protect these responders. If there are more resources available, more responders may be available throughout the domain of the implementation. Repository redundancy and network distance are both rated average.

12. Simple Certificate Validation Protocol (SCVP)

SCVP is still an IETF draft specification; however, the draft provides enough information to evaluate the proposed scheme. SCVP is designed to reduce the burden on constructing and validating certification paths for end user applications while offloading certificate handling to a trusted SCVP server. The server can provide certificate status or

can be used to build the chain of trust from the relying party to an authoritative server. The SCVP is a responder method that provides information culled from other sources. As such, the provided information may be as current as information provided by external sources. This flexibility is evaluated as above average.

a. Data Structure Response Size

It is not clear from the IETF specifications as to how large the response will be; however, it is clear that the SCVP servers should be able to cache revocation information and reply to the replying party with validation information. This response will ultimately be a response to a particular validation request and be relatively small in size. This is evaluated as above average.

b. Bandwidth Profile

SCVP severs may be implemented as either trusted or untrusted. Similar to an OCSP response, the lightweight nature of the response makes it inherently tolerant of low bandwidth environments. Also similar to OCSP, SCVP may be less tolerant of intermittent bandwidth environments. Unlike OCSP, SCVP has twenty-two separate responses that might be used to indicate to the relying party both the status of the server and the connection. It is evaluated as above average for bandwidth throughput and average for intermittent bandwidth environments.

c. Response Generation Latency

The SCVP server must generate a response that will be interpreted by the relying party application. Depending on whether the implementation is trusted or untrusted, the SCVP server may process both signed and unsigned requests and reply with either a signed or unsigned response. The work required to interpret signed requests and formulate signed responses is not trivial but may be offset by increased SCVP servers throughout the domain and should be similar to that of OCSP. Response generation latency is rated as average.

d. Proximity

The SCVP implementation requires handoff of requests and forwarding of data. The query forwarding and handoff is rated as below average. Network distance and repository redundancy are both dependent on the type of implementation. A best case implementation would have numerous SCVP servers within the domain and

numerous distributed redundant repositories from which to draw. While information redundancy is rated as above average, the degree of query forwarding and handoff are rated below average. SCVP works as an intermediary mechanism to CRL's, similar to OCSP responders, but it may also act as an intermediary to an OCSP responder. This adds another layer of complexity between the relying party and the information provided by the CA.

13. Micro-CRL Mechanism

The Micro-CRL mechanism is a significant departure from the mechanisms discussed thus far because the onus of revocation checking is shifted to the sending party. This scheme provides certificate validation information that is sent with the certificate presented to the relying party; this allows the relying party to make immediate trust decisions upon receiving the certificate. In a tactical environment there may be utility for an implementation wherein the sending party knows that the relying party can act on the sent information upon reception…not after some indeterministic amount of time due to un-reliable revocation checking performed by the relying party. The key link in the trust model is that of the CA; because the relying party does not interface with the CA, he has no information as to whether the CA certificate used to sign the revocation information is still valid. The occurrence of a compromised CA certificate is so rare as to be considered insignificant. The Micro-CRL mechanism fits within the X.509 structure and is not a large departure from the current architecture. Because this implementation has not been fully explored the previously discussed limitation regarding the lack of communication with the CA prevents it from being a serious choice for the DoD, however, it should not limit further research in this area.

a. Currency

The currency of this mechanism will be the same as a full CRL. While there is no formal draft to support this method, it resembles currently fielded PKI implementations in that at a specified periodicity the CA will produce a large set of signed Micro-CRLs that are comprised of a signed upper and lower boundary of revoked certificates. This is similar to producing only the leaves from a Merkle tree and then signing them. The CA would make these available for download on distributed servers from within the domain. The sender would then download only the relevant Micro-CRL

from the repository and utilize it. The implications are that the currency is the same as any other periodic publication mechanism and is as defined by the CP and CPS, which in practice, is generally twenty-four hours. The data remains valid until the *NextUpdate* time; at which time a new Micro-CRL must be produced by the CA. This meets the criteria for medium currency and a rating of average.

b. Data Structure Response Size

While this method has not been formalized, it is supposed that the data response structure size would be on the order of a few bytes; enough information to relay the upper and lower bounds of valid certificates and provide a digital signature. This might be similar in size to an OCSP response, that is, very small. Data structure response size is evaluated as above average.

c. Bandwidth Profile

This implementation is tolerant of both low bandwidth and intermittent throughput environments. All data required to process the transaction is sent with the message. The bandwidth profile for the relying party is above average while the bandwidth profile for the sending party is above average; the sending party must accept the onus of downloading the Micro-CRL. The net effect to the entire PKI is the same; the bandwidth requirements have been shifted between parties. The Micro-CRL implementation would best be suited for a tactical environment in which the relying party may experience low bandwidth, utilize devices with limited processing power, or be required to make immediate trust decisions based upon the presented information. It is easy to imagine that this type of mechanism might be of particular merit to a relying party on the battlefield, submarines, or Special Forces personnel. The transaction may be completed in any environment in which the signed message may be received by the relying party. If the bandwidth is such that a message cannot be received by the relying party, there is no requirement for revocation checking. Bandwidth profile is evaluated as above average.

d. Response Generation Latency

Response generation latency is the onus of the sending party. The computational power to complete this transaction is relatively small and should be available not only on robust computers but also on handheld devices. Response generation latency is evaluated as above average.

14. Novomodo

Novomodo works with the X.509 v3 standard certificates but end user applications would be required to be built for this scheme to be used within a DoD PKI. A security concern not directly measured by the metrics is the need for security of hashed data. While most PKIs require the security of the public key, the Novomodo process requires that intermediate hashed data be kept confidential. If this data is leaked, the entirety of the PKI is compromised; the same can be said of a traditional PKI, if the key is compromised the entirety of the PKI is compromised.

a. Currency

As with many of the other systems discussed, the currency of Novomodo is limited only by the rule sets defined in the CP and CPS. The limitation on how often a Novomodo proof might be released is defined by the life of the certificate, the frequency of updates, and the processing work available to support the PKI. As previously discussed in the section on Novomodo, an end hash must be computed that becomes the validity target. In the example this was X365 where the life of the certificate was one year and the CA updated the hash daily (e.g., X*today* = hash(X*yesterday*). The number of hashes which must be computed for each certificate is equal to the number of days in the life of the certificate multiplied by the number of proofs released each day; X(number days * number of times per day a proof is released). The 20 bit values from X(first proof) to X(last proof-1) must be kept secret until they are released as proof; once released the hashed values require no secrecy. It is possible that in a DoD PKI the ability to keep these small 20 bit values safe can be guaranteed. For example, the 2003 end strength for Active Duty Military was approximately 1,500,000 personnel.[57] If each member had a five year certificate and a proof was released every eight hours, the number of values that would have to be safeguarded is equal to 1,500,000 persons * 5

[57] Military Personnel Statistics, [http://www.dior.whs.mil]. Accessed Nov 2003.

years * 365 days * 3 updates/day = 8,212,500,000 twenty bit values, or roughly 19GB of data would have to be protected. The amount of processing work to formulate these hashes would be substantial but feasible. This attribute is evaluated as above average.

Another method for implementing Novomodo might be in an environment where secure storage of the intermediate users is not feasible. By computing the validity target and revocation target (the final hash value) and then destroying all intermediate hashes, the CA could maintain only the initial values Y0 and X1 and compute the new validity targets prior to the scheduled *NextUpdate* time. While this would require some computational power, processing a hash for 1,500,000 DoD users, this is not outside the realm of possibility. While the original paper describing Novomodo method specified proof release times, it would be feasible to instead configure an online "in demand" method of release. In this way, proofs would only be computed when requested and would significantly reduce the amount of computational power required to produce these proofs.

b. Data Structure Response Size

Depending on the implementation, the data structure response size can be as small as twenty bits of data accompanied by several bits of administrative information. The Novomodo method might be implemented via an online responder system or as a periodic publication that could be emailed, downloaded, or made available on the web. An additional note worth making is that while the Novomodo method requires secrecy of the original key values and any non-released interim proofs, data released to LDAP servers throughout the domain does not require protection nor is a key required on any of the LDAP servers; the mathematical test will determine if the information has been tampered with. Data response size is evaluated as above average.

c. Bandwidth Profile

In the previous example, for a domain of 1,500,000 users, a release of proofs for 1,500,000 users is equal to approximately 3.6 Megabytes of information (1,500,000 users * 20 bits = 30,000,000 bits or 3.5762 Megabytes of data). A download or email method for this information would not be tolerant of low bandwidth or intermittent connectivity, however, an implementation using an online responder would only be required to provide a twenty bit response. In order to provide tolerance to

intermittent connectivity, the entirety of the list would be required to be cached and would therefore not be tolerant of low bandwidth environments. These attributes of bandwidth and network connectivity are inversely related; each is rated as above average but with the caveat that both cannot be maximized at the same time.

d. Response Generation Latency

There are two components to the response generation latency. The first is the mathematical proof that is computed by the CA in formulating the validity and revocation targets. The Novomodo method requires that the end hash--the validity target--and all intermediate leading up to the validity target be computed before the certificate is placed into circulation. By computing these values ahead of time the amount of processing required to produce these proofs is completed before the first proof must be released and are therefore of no consequence to the total time to generate a response. The second component occurs when the relying party receives the response and must generate the mathematical hash required to test the validity of the supplied proof. This hashing and checking function is lightweight enough to require only minimal work and should be able to be completed on the smallest processor, to include most modern handheld devices. This metric is evaluated as above average.

e. Proximity

The network distance and information location redundancy in this proposed model could be maximized depending on the type of implementation. It is relatively trivial to minimize network distance by placing responders or cached copies of the most recent proofs within the network domain. For the Novomodo method, there is no requirement to keep any secret keys on the distributed LDAP or caching servers. If an attacker were to break into a LDAP and change the proof, the new proof would neither match the validity or revocation targets. Information location redundancy may be optimized by increasing the number of responders with copies of the cache throughout the domain. Both of the metrics are evaluated as above average. It is envisioned that there would be no requirement for query forwarding or handoff.

15. MiniCRL

The miniCRL uses the X.509 structure but does not strictly conform to the X.509 v3 CRL; in order to be used within the DoD PKI a new infrastructure would need to be implemented and an end user application would need to be built. In terms of security attribute rating, the miniCRL is similar to the complete CRL with the exception of the data structure response size. The size may be reduced by an estimated 30:1 ratio (reduction to roughly 3.3 percent the original size); however, some of this reduction is due in part to compression that could also be applied to a normal CRL. The compression achieved solely from redesigning the CRL structure is estimated at an 18:1 ratio (reduction to almost 5.5 percent the original size). This indicates that the data structure response size is smaller but the compression gains should not be used as a reference point because similar compression gains could be achieved by the full CRL. The MiniCRL has the smallest data response size of the periodic publication methods. Data structure response size is evaluated as average.

16. Short-Lived OID Attribute Certificates

This method uses short-lived certificates that expire soon after receipt of the relying party; there is no method of certificate revocation because the certificate life is in effect the revocation point. It does not fit within the ranking system that was built for certificate revocation mechanisms. This method is best placed in a closed environment where sufficient computational power is available to create certificates on a routine basis and sufficient bandwidth if available to deliver these certificates in a timely manner to ensure their receipt before expiration. This method is often seen in private PKI systems such as banking or financial institutions.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSIONS

A. SECURITY STRENGTH ORDERING CONCLUSIONS

1. Security Ordering Strength of Revocation Schemes

The proposed "best case" revocation method is the highest ranked of all examined methods. This is not a surprising outcome; the best case method is a hybrid mechanism designed to minimize all unacceptable attributes while maximizing those that strengthen the implementation in terms of security attributes. This process of describing the "best case" method was undertaken with little regard for the dollar amount, infrastructure, or manpower cost associated with achieving these strengths. All of the implementations have inherent strengths and are able to maximize others by fielding additional infrastructure in the form of servers for computational power, LDAP directories, or other types of equipment and resources. The increased cost to achieve these benefits; however, might exceed the potential increase in security.

Figure 17 is a graphical depiction of all plausible implementations of certificate revocation management ordered by security strength based upon the preceding examination of security attributes.

			Bandwidth Profile			Proximity		
● = Above Average O = Average ◠ = Below Average / = Does not apply Implementation	Currency	Data Structure Response Size	Mean Bandwidth / Throughput	Steady State vs Intermittent	Response Generation Latency	Network Distance	Degree of Forwarding / Handoff	Repository Redundancy
Proposed "Best Case" Revocation Mechanism	●	●	●	●	●	●	●	●
NOVOMODO (aka Mikali method)	●	●	●	●	●	●	/	●
Micro-CRL	O	●	●	●	●	/	/	●
Online Certificate Status Protocol (OCSP)	●	●	●	O	O	●	/	●
Simple Certificate Validation Protocol (SCVP)	●	●	●	O	O	●	◠	●
Certificate Revocation Trees (CRTs)	●	●	●	●	◠	O	/	O
Sliding Window Delta CR	O	●	●	◠	◠	O	O	O
Indirect Delta CRLs	O	●	◠	◠	●	O	O	O
Mini CRL	O	O	◠	◠	◠	O	/	O
CRL Distribution Points or Partitioned CRLs	O	O	◠	◠	●	O	O	O
Redirect CRLs and Enhanced CRL Distribution Point	O	O	◠	◠	●	O	O	O
Complete CRL	O	◠	◠	◠	●	O	/	O

Figure 17. Revocation Method Rankings.

The Novomodo method was a close second to the best case model. The rationale for not having it tie the "best case" model is that there are no currently known fielded Novomodo systems. While the Novomodo model is X.509 version 3 compliant, the infrastructure to implement this model is not currently in place. The most significant weakness not measured by the metrics is the need to keep a large number of proofs safeguarded until the designated release time; it has been shown that this weakness may also be negated depending on implementation of the system. Novomodo is a very strong, scalable mechanism for certificate revocation that deserves more research.

In third place in the rankings, the MICRO-CRL has several strengths that should not be overlooked. By assuming responsibility for revocation checking, the sender is confident that immediate action may be taken upon receipt of the certificate by the relying part, not after some indeterministic amount of time due to unreliable revocation checking. In a tactical environment, this ability to act in near real time may be of great value.

In fourth place, the implementation which most closely resembles the "best case" method is that of OCSP. OCSP encompasses many of the same building blocks used in the "best case" scenario. The significant differences were a result of the "best case" method being a hybrid that capitalized on the strengths of OCSP and the redundancy of more traditional models. A well thought out robust OCSP implementation will have many of the same attributes as the "best case" model, a generic installation of OCSP will not.

While relatively high in the rankings, SCVP is more of a validation authority than a different type of revocation scheme because this method acts an intermediary for any number of revocation mechanisms and provides that information to the relying party while performing the additional tasks of certificate path validation. It can be argued that OCSP also acts as a validation authority to some extent; however, SCVP adds another layer of complexity with the possibility of relaying OCSP responses that are in turn relaying information from CRL's obtained from the CA.

2. Implications for DoD PKI

The final research question also serves as a summary for this thesis: Of all the plausible implementations of certificate revocation management, which are the best candidates with regard to interoperability for DoD use? The answer depends on resources and willingness to change on the part of stakeholders. The information presented in this thesis makes it clear that Novomodo and MiniCRLs warrant further research and discussion as methods that should be used for a secure and scalable PKI. Each of these methods has strengths relative to the environment in which they might be placed.

A September 2003 Navy message indicated that well over one million PKI enabled access cards had been issued to DON personnel; this message also mandated that all commands achieve compliance with the OCT 2003 DoD PKI milestones.[58] There has been a significant investment in time, money, and infrastructure in the X.509 standard. It is possible that that the Novomodo or Micro-CRL methods could have been optimized, their weaknesses reduced, and implemented into the DoD infrastructure. With the amount of progress that has been made in securing the X.509 infrastructure, this is now unlikely. Which then, of the remaining X.509 compliant PKI methods should be considered for implementation by DoD? A strong case may be made for a hybrid OCSP implementation; the end result of such an implementation should resemble the "best case" proposal. This implementation would confer a greater sense of security to the relying party while the highly redundant mixed architecture would reliably deliver validation information, or state that the information is unobtainable and give the relying party trust options from which to choose.

3. Further Research

Validation Authority. Several companies have begun to offer Validation Authority services. Instead of the standard certificate services (registration, validation, revocation, etc.), this service delivers aggregated validation information to the relying party. These services are generally negotiated through the organization to which the relying party belongs. The company providing the service may or may not be the CA that is generating the revocation information. An example of this product is the

[58] DON CIO WASHINGTON DC RMG 151407ZSEP03.

Valicert Validation Authority (VVA) Service. As depicted in Figures 18 and 19, the service draws information from a variety of CA sources and makes it available to the relying party.

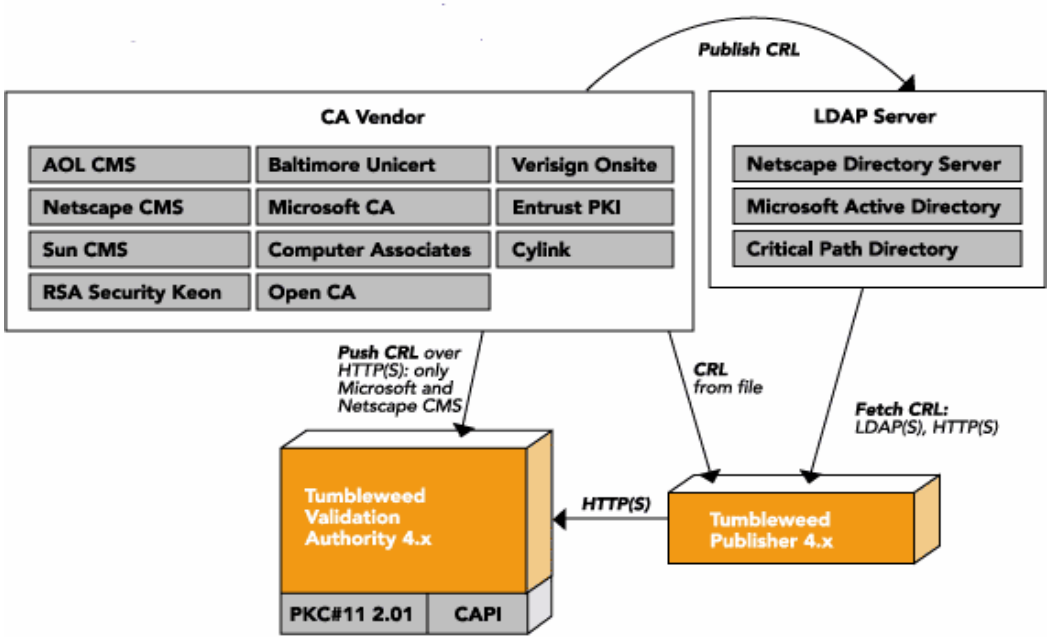

Figure 18. Valicert Validation Authority Architecture, Server Architecture[59].

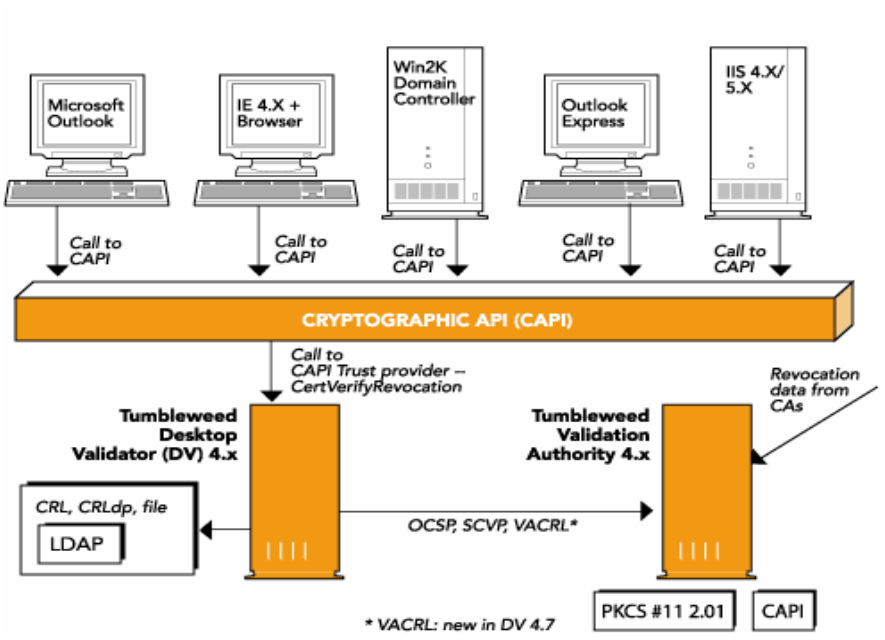

Figure 19. ValicertValidation Authority Architecture, Client Infrastructure[60].

[59] [http://www.valicert.com/products/validation_authority.html]. Accessed Nov 2003.

[60] Ibid.

The VVA presents the relying party with three choices of validity checking: Fetch and cache the CRL and use it as a basis for checking revocation status (CRL/CRL DP), query an OCSP responder, or allow SCVP to delegate chain building and validation to the Validation Authority.[61]

Although developed independently, the Validation Authority and "best case" models share similarities in that each is a distributed model that offers choices to the relying party and allows both the mechanisms and the relying party to follow several chains of trust. The OCSP Conops model might be thought of as the precursor to Validation Authority. Further research is warranted in testing hybrid models that make use of several chains of trust, offer redundant distributed networks, and allow the user to make informed decisions on the information available.

Hybrid Novomodo and Mini-CRL Method. The "best case" scenario was an attempt to create an "ideal" baseline against which actual methods could be compared. It is not reasonable to believe that the assets would be allocated to field such an implementation. A new model is offered in this thesis for consideration and further study; this is the Stellerjay model, named for the common and widespread bird resident in coastal and montane coniferous forests throughout California. The Stellerjay model is a hybrid in which the onus of providing certificate revocation is on the sending party, and Novomodo proofs are substituted for the signed Mini-CRL data. When sending a signed document, the sending/signing party provides verification to the relying party that the private signing key he used is still valid. If the relying party receives signed data without this proof, she should not trust either the claimed source or the integrity of that data.

During each update period, the CA provides the new validity proofs to public LDAP directories throughout the domain. Timestamps are not required because each Novomodo proof is inherently tied to a specific validity period. In order to provide the proof to the relying party, the sending party connects to an online responder that looks up the identity of the sending party and responds with the appropriate proof. The sending party includes this proof with the signed data to the relying party.

---

[61] Ibid.

One update's worth of Novomodo proofs constitutes a sizeable amount of data (3.5762 Megabytes for an environment of 1,500,000 users) which must be sent to LDAP servers. However, not all of this information need reside on the all LDAPs; because the relying party is retrieving the information before sending it, proofs segmented by organization may allow for smaller transmissions to specific servers. For example, the June 2003 end strength was 334,264 Naval Personnel. Had the proofs for these sending parties been sent to a server within their organization; the total amount of data to be transmitted would have been less than one megabyte of information to transmit and store (334,264 users * 20 bits = 0.796 Megabytes). Further segmentation within the organization would provide continued reductions in the amount of data required to be sent to particular subsets within the domain.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California